# Competition and Loss of Efficiency: From Electricity Markets to Pollution Control



by

Lionel J. Kluberg

Submitted to the Sloan School of Management
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011



Author ........................................................
Sloan School of Management
May 20, 2011

Certified by ...................................................
Georgia Perakis
William F. Pounds Professor of Operations Research
Sloan School of Management
Thesis Supervisor

Accepted by ...................................................
Dimitris Bertsimas
Boeing Professor of Operations Research
Co-Director, Operations Research Center

# Competition and Loss of Efficiency: From Electricity Markets to Pollution Control

by

Lionel J. Kluberg

Submitted to the Sloan School of Management
on May 20, 2011, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

## Abstract

The thesis investigates the costs and benefits of free competition as opposed to central regulation to coordinate the incentives of various participants in a market. The overarching goal of the thesis is to decide whether deregulated competition is beneficial for society, or more precisely, in which context and under what market structure and what conditions deregulation is beneficial. We consider oligopolistic markets in which a few suppliers with significant market power compete to supply differentiated substitute goods. In such markets, competition is modeled through the game theoretic concept of Nash equilibrium. The thesis compares the Nash equilibrium competitive outcome of these markets with the regulated situation in which a central authority coordinates the decision of the market participants to optimize social welfare. The thesis analyzes the impact of deregulation for producers, for consumers and for society as a whole. The thesis begins with a general quantity (Cournot) oligopolistic market where each producer faces independent production constraints. We then study how a company with multiple subsidiaries can reduce its global energy consumption in a decentralized manner while ensuring that the subsidiaries adopt a globally optimal behavior. We finally propose a new model of supply function competition for electricity markets and show how the number of competing generators and the electrical network constraints affect the performance of deregulation.

Thesis Supervisor: Georgia Perakis
Title: William F. Pounds Professor of Operations Research
Sloan School of Management

# Acknowledgments

To begin, I would like to thank my advisor Georgia Perakis for her mentorship throughout my PhD. The breadth of her knowledge as well as her sincere care and kindness have been infallible driving forces behind my work. Georgia not only gave me my first research topic, she also provided me with unlimited trust and support to dig into mathematical problems and explore new research topics.

Georgia deeply cares about the professional fulfillment of her students. She sent me to many conferences to present my work and perfect my communication skills and encouraged me to TA a number of classes including a core MBA course. She was infinitely generous with her time and advice in difficult times. Georgia is a great researcher and a skilled manager. Her student dinners punctuated my life at MIT with happy memories.

I would also like to thank the members of my committee Dimitris Bertsimas and Retsef Levi for their feedback and valuable comments on my work. On the academic side still, I learnt from a number of wonderful teachers at MIT. I would especially like to thank David Gamarnik, Retsef Levi, Gustavo Manso, Asu Ozdaglar, Pablo Parillo and Andreas Schulz for all they taught me. I will always be indebted to Frederic Bonnans and Stephane Gaubert from Ecole Polytechnique for introducing me to Operations Research and directing me to MIT.

The Operations Research Center has been my home at MIT for the last five years. A great thanks to Cindy Barnhart, Dimitris Bertsimas and Patrick Jaillet for their dedication at the head of the Center. My sincere appreciation also goes to Laura Rose and Andrew Carvalho for making sure the center runs smoothly and the students have everything they need to work.

Life in the United States would not have been the same without the presence of my friends. I interacted with many wonderful colleagues and friends at the Operations Research Center. In particular, I enjoyed working and spending time with Adrian Becker, David Goldberg, Dan Iancu, Thibault Leguen, Ilan and Ruben Lobel, Joline

Uichanco, Sun Wei and Gareth Williams. I would also like to thank my roommates Michael Bikard, David Opolon, Martin Quinteros, Alexander Rikun and Carine Simon who had the courage to live with me over the years. A special thanks to Alexis Turjman, Maxime Cohen, Sivie Naimer and the Zyto family for always being there for me to have fun and support me.

Throughout my PhD, I have also been priviledged to work with some extremely talented people on non research related projects. I would like to thank Alex Hagiage who cofounded MyeShoe with me, the GSC 2008 Career Fair committee, MIT Ventureships and the Consulting Club at MIT for enriching my life with interesting and fun experiences. A special thanks to Highvista Strategies for believing in me, for providing me with a job and for being such an incredible environment to work in.

Another key aspect of my life in Cambridge has been my involvement in the local jewish community. I cannot thank enough Rabbi Ganz and his wife, the Harvard Chabad and MIT Hillel for the community they help build, for the fabulous meals I have shared with them and for their insightful teachings.

Last but not least, I am extremely grateful to my father and my sister for encouraging me to come to MIT. Going through the ups and downs of PhD life at the same time as my sister has been a wonderful experience. I would also like to express my eternal gratitude to the Cohen-Arazi, the Ejnes and the Touati families who have had an essential role in making me who I am today. Finally, I dedicate this thesis to my mother whose memories will accompany me forever.

# Contents


**1 Introduction** **15**

**2 Generalized quantity competition for multiple products and loss of efficiency** **17**

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.1.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . 17

2.1.2 Contributions . . . . . . . . . . . . . . . . . . . . . . . . 20

2.2 Model and assumptions . . . . . . . . . . . . . . . . . . . . . . . . 21

2.3 Loss of total surplus and profit under production capacities . . . . . . 28

2.3.1 Loss of total surplus . . . . . . . . . . . . . . . . . . . . . 28

2.3.2 Loss of firms profit . . . . . . . . . . . . . . . . . . . . . 35

2.4 Loss of total surplus and profit under general constraint sets . . . . . 44

2.4.1 Loss of total surplus . . . . . . . . . . . . . . . . . . . . . 45

2.4.2 Loss of firms profit . . . . . . . . . . . . . . . . . . . . . 52

2.5 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

**3 Enforcing a joint energy consumption target on companies with multiple subsidiaries: a decentralized incentive mechanism and its social impact** **59**

3.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 59

3.1.1 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . 59

3.1.2 Literature review . . . . . . . . . . . . . . . . . . . . . . 62

3.1.3 Main contributions and chapter outline . . . 66
3.2 Model Description . . . 68
3.2.1 Cournot competition under a joint constraint . . . 72
3.2.2 Reward-based incentive mechanism . . . 74
3.2.3 Equivalence of the models . . . 76
3.2.4 Company and social objective . . . 78
3.3 The Need for Coordination . . . 79
3.4 Many Symmetric Subsidiaries . . . 83
3.4.1 Assumptions and closed-form solutions . . . 83
3.4.2 Loss of profit and welfare . . . 84
3.4.3 Underlying production strategies . . . 92
3.5 Many Subsidiaries with Symmetric Price Potentials . . . 93
3.5.1 Closed-form solutions . . . 94
3.5.2 Loss of profit and welfare . . . 94
3.5.3 Underlying production strategies . . . 97
3.6 Extension to General Price Potentials . . . 99
3.7 Conclusions . . . 104

**4 Loss of efficiency in deregulated electricity markets: a supply function equilibrium approach 107**

4.1 Introduction . . . 107
4.1.1 Motivation . . . 107
4.1.2 Literature review . . . 110
4.1.3 Contributions and Outline . . . 113
4.2 The supply function equilibrium (SFE) model . . . 115
4.3 Unconstrained symmetric generators . . . 124
4.3.1 The case of a monopolistic generator . . . 126
4.3.2 The case of two generators . . . 128
4.3.3 The case of many generators . . . 130

4.4 Unconstrained asymmetric generators . . . . . . . . . . . . . . . . . . . 134
4.4.1 The model . . . . . . . . . . . . . . . . . . . . . . . . . . . 134
4.4.2 Market insights . . . . . . . . . . . . . . . . . . . . . . . . 136
4.5 Performance under general network constraints . . . . . . . . . . . . . 146
4.5.1 Simulation method . . . . . . . . . . . . . . . . . . . . . . . 147
4.5.2 The 30-bus system . . . . . . . . . . . . . . . . . . . . . . . 150
4.5.3 The 5-bus system . . . . . . . . . . . . . . . . . . . . . . . . 154
4.6 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 158

**Conclusions** **161**

**A Proofs for Chapter 2** **163**
A.1 Proof of Theorem 2.1 . . . . . . . . . . . . . . . . . . . . . . . . . 163
A.2 Proof of Step 1 for Theorem 2.3 . . . . . . . . . . . . . . . . . . . . 168
A.3 Calculations for Theorem 2.4 . . . . . . . . . . . . . . . . . . . . . 169
A.4 Proof of Lemma B.1 . . . . . . . . . . . . . . . . . . . . . . . . . . 171
A.5 Derivation of oligopoly variational inequality . . . . . . . . . . . . . 172

**B Proofs for Chapter 3** **175**
B.1 Calculations and proofs for Section 3.3 . . . . . . . . . . . . . . . . . 175
B.1.1 KKT conditions . . . . . . . . . . . . . . . . . . . . . . . . . 175
B.1.2 Proofs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 176
B.1.3 Calculations for the duopoly case . . . . . . . . . . . . . . . . 182
B.2 Calculations and proofs for Section 3.4 . . . . . . . . . . . . . . . . . 184
B.2.1 Closed-form solutions . . . . . . . . . . . . . . . . . . . . . . 184
B.2.2 Proofs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 185
B.3 Calculations and Proofs for Section 3.5 . . . . . . . . . . . . . . . . . 198
B.3.1 Closed-form solutions . . . . . . . . . . . . . . . . . . . . . . 198
B.3.2 Proofs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 198
B.4 Calculations and Proofs for Section 3.6 . . . . . . . . . . . . . . . . . 205

**C Proofs for Chapter 4** **211**

C.1 Derivations for unconstrained symmetric generators . . . . . . . . . . . 211

C.1.1 Supply function equilibrium derivation . . . . . . . . . . . . . . 211

C.1.2 Calculations of surplus, profit and welfare ratios . . . . . . . . . 214

C.2 Derivations for unconstrained asymmetric generators . . . . . . . . . . 216

# List of Figures


2-1 Numerical simulation of the profit ratio $\Pi(OP)/\Pi(MP)$ for general unconstrained markets (in blue) and theoretical bound for uniform markets (in red). . . . . . . . . . . . . . . . . . . . . . . . . . . 44

2-2 Numerical simulation of $\Pi(OP)/\Pi(MP)$ for general constrained (in purple) and unconstrained (in blue) markets. . . . . . . . . . . . . . . 45

2-3 Bounds on the profit ratio $\Pi(OP)/\Pi(MP)$ for markets with simple capacity constraints (green) and general convex constraints (red). . . 53

3-1 Stackelberg subsidy game: $w$'s are set parameters, $\mu$ and $\mathbf{d}$ are the decision variables. . . . . . . . . . . . . . . . . . . . . . . . . . . 75

3-2 Profit loss $(1 - \Phi_1(r, n))$ for free competition under unlimited resources 86

3-3 Maximum profit loss $(1 - \Phi_2(r, n))$ for free competition under limited resources . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87

3-4 Profit loss for the worst equilibrium . . . . . . . . . . . . . . . . . . . 87

3-5 Effect of the joint constraint on the loss of profit . . . . . . . . . . . . 88

3-6 Maximum loss of social welfare $(1 - \Phi_4(r, n))$ under free competition with limited resources . . . . . . . . . . . . . . . . . . . . . . . . . . 90

3-7 Effect of the joint constraint on the loss of social welfare . . . . . . . 91

3-8 Loss of profit for the uniform Nash equilibrium . . . . . . . . . . . . . 96

3-9 Loss of social welfare for the uniform Nash equilibrium: n=2 . . . . . 97

3-10 Loss of social welfare for the uniform Nash equilibrium: n=5 . . . . . 98

3-11 Loss of social welfare for the uniform Nash equilibrium: n=10 . . . . 98

4-1 Electricity market diagram. . . . . . . . . . . . . . . . . . . . . . . . . 116
4-2 Welfare, surplus and profit ratios for a monopolistic generator as functions of the unit-less cost parameter $z$. . . . . . . . . . . . . . . . . . . 129
4-3 Welfare, surplus and profit ratios for two symmetric generators as functions of the unit-less cost parameter $z$. . . . . . . . . . . . . . . . . . . 130
4-4 Profit ratios for $n$ symmetric generators ($n$ ranging from 2 to 7) as functions of the unit-less cost parameter $z$. . . . . . . . . . . . . . . . . 131
4-5 Welfare ratios for $n$ symmetric generators ($n$ ranging from 2 to 7) as functions of the unit-less cost parameter $z$. . . . . . . . . . . . . . . . . 133
4-6 Consumer surplus ratios for $n$ symmetric generators ($n$ ranging from 2 to 7) as functions of the unit-less cost parameter $z$. . . . . . . . . . 134
4-7 Consumer surplus ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds. . . . . . . . . . . . . . . . . . . . . . . . . . . . 141
4-8 Consumer surplus ratio for a duopoly ($z_1 = 0.5$, $z_2$ varies) with simplified lower bound. . . . . . . . . . . . . . . . . . . . . . . . . . . . 141
4-9 Profit ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds (4.16). . . . . . . . . . . . . . . . . . . . . . . . . . . . 143
4-10 Profit ratio for a duopoly ($z_1 = 1$, $z_2$ varies) with simplified upper bound (4.17). . . . . . . . . . . . . . . . . . . . . . . . . . . . 144
4-11 Social welfare ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds. . . . . . . . . . . . . . . . . . . . . . . . . . . . 145
4-12 Social welfare ratio for a duopoly ($z_1 = 1$, $z_2$ varies) with simplified upper and lower bounds. . . . . . . . . . . . . . . . . . . . . . . . . 146
4-13 30-bus power system test case. . . . . . . . . . . . . . . . . . . . . . . 151
4-14 5-bus power system diagram. . . . . . . . . . . . . . . . . . . . . . . . 155

# List of Tables


4.1 IEEE 30-bus system market data. . . . . . . . . . . . . . . . . . . . . . 150

4.2 30-bus system generating cost data. . . . . . . . . . . . . . . . . . . . . 151

4.3 30-bus system supply function equilibrium output. . . . . . . . . . . . . 152

4.4 30-bus system Cournot output. . . . . . . . . . . . . . . . . . . . . . . 152

4.5 30-bus system centrally coordinated output. . . . . . . . . . . . . . . . 153

4.6 5-bus distribution factor matrix. . . . . . . . . . . . . . . . . . . . . . 155

4.7 5-bus line thermal capacities. . . . . . . . . . . . . . . . . . . . . . . 155

4.8 5-bus system generating cost data. . . . . . . . . . . . . . . . . . . . . 156

4.9 5-bus supply function equilibrium dispatch. . . . . . . . . . . . . . . . 156

4.10 5-bus Cournot dispatch. . . . . . . . . . . . . . . . . . . . . . . . . . 157

4.11 5-bus centrally coordinated dispatch. . . . . . . . . . . . . . . . . . . 158

# Chapter 1

# Introduction

This thesis investigates the costs and benefits of free competition as opposed to central regulation to coordinate the incentives of various participants in a market. The overarching goal of the thesis is to decide whether deregulated competition is beneficial for society, or more precisely, in which context and under what market structure and what conditions deregulation is beneficial. The market setting we study in this thesis consists of i) a small number of producers whose goal is to optimize their own profit and ii) price taking consumers. The thesis considers oligopolistic markets, i.e. only a few producers participate in the market and the decisions of each producer affect the overall market price. Since each producer's decisions affect the profit of its competitors, the market outcome is modeled through the game theoretic concept of Nash equilibrium. That is, a market equilibrium is a set of producers' choices such that no producer wants to modify its choice unilaterally. The thesis analyzes how the prices set in the market and the quantities sold are affected by competition. We also look at the impact of free competition on producers' profit, consumers surplus (benefit) and the welfare of society as a whole. For each of these measures, the thesis estimates the ratio of the performance of free competition over that of a regulated market where prices and quantities are fixed by a central authority which seeks to maximize social welfare. The thesis evaluates this ratio, often referred to as the *price of anarchy* in the economics literature, as a function of market characteristics such as the number of competing generators, the number of products sold and the intensity of competition.

We look at the impact of deregulation in three contexts. We begin in Chapter 2 with a general quantity (Cournot) oligopolistic market. A few firms sell multiple differentiated products and face production constraints. This setting models competition between the US wireless carriers for example. Each carrier offers multiple products in the form of different phone plans. Each carrier faces service constraints due to the capacity of the wireless network. In this market, carriers compete for market share (quantity) much more than they compete through prices.

Chapter 3 considers a company with multiple subsidiaries acting independently which attempts to reduce its global energy consumption level. Each subsidiary still seeks to maximize its selfish profit but the parent company cares about the global profit as well. Whereas in Chapter 3, each producer faces independent constraints on its production set, the energy target of the parent company is a constraint that ties the subsidiaries together. This adds another level of complexity to the analysis and radically changes the performance of free competition.

Chapter 4 proposes a new model of supply function competition for electricity markets. Several generators bid electricity to a system operator by submitting a supply function, i.e. the quantity they are willing to produce as a function of electricity price. After receiving these bids, the system operator dispatches production among the generators to obtain the best solution for society that respects the generators' bids. We compare the performance of such a deregulated market with that of a traditional regulated utility system where the system controls the generators, knows their production costs, and optimizes social welfare (without receiving bids from the generators). The chapter studies the performance of deregulation as a function of the number of generators participating in the market and the constraints imposed by the electrical network.

Finally, the last Chapter provides some conclusions.

# Chapter 2

# Generalized quantity competition for multiple products and loss of efficiency

## 2.1 Introduction

### 2.1.1 Motivation

Traditional economic theory states that in a "perfect" market, free competition between firms is optimal for society. One key assumption behind a perfect market, is that the number of competing firms is very large and that each firm is individually too small to affect the market. Since the papers of Cournot [33] and Bertrand [17], economists have studied imperfect markets where only "a few" firms compete. The outcome of these so-called oligopoly markets is represented as the Nash equilibrium [92] of a single shot static game between the firms. While some papers focus exclusively on establishing existence and/or uniqueness of an equilibrium ([124], [111], [89], [8], [37], [93], [134]), most papers are also interested in the performance of these markets. Literature on the subject is too vast to be extensively listed, but a few main approaches emerge. Some papers study directly the relation between market structure and equilibrium price, production cost, or profit ([86], [34], [117], [120], [40], [118], [49]). However, these papers do not study social welfare. Other papers introduce concentration indices, one-dimensional measures aimed at summarizing the market structure, and show the relation between these indices and profit or welfare ([5], [122], [36], [44], [35]). As Tirole argues though, in Chap 5 of [128], no index is perfectly correlated with either profit or welfare when firms are asymmetric. Very few papers, on the other hand, study directly the relation between market structure and social welfare. From a policy perspective, social welfare is arguably the quantity that matters most: a policy should be judged on its ability to improve welfare. The goal of this chapter is to study the relation between market structure and the loss of welfare in an oligopoly market.

In other contexts, there is a large literature studying the loss of efficiency resulting from competition. Koutsoupias and Papadimitriou [77] introduced the concept of *price of anarchy* to denote the ratio between the worst case Nash equilibrium resulting from competition (see [92]) and the coordinated solution. In the field of transportation, many papers study the loss of efficiency (e.g. [31], [102], [115], [114]). The game there involves non-atomic (infinitely small) players. Applications to supply chain management have also been studied recently ([28], [38], [103], [26]). While in these settings the goal of players is to minimize costs, our research focuses on profit maximization. Results do not transfer as some assumptions differ.

In the context of oligopolistic competition, two models of markets have emerged: price (Bertrand) competition and quantity (Cournot) competition. In the former, firms compete by setting prices letting consumers decide what quantity they buy, while in the latter, firms decide production quantities and set prices to clear the market. The loss of efficiency under price competition has been extensively studied by Farahat and Perakis ([46], [47], [48]). This chapter, in contrast, deals with quantity competition. Unfortunately, the results and insights they obtain as well as the proof techniques they employ do not transfer from the price to the quantity competition setting (see [46] for a discussion on how the two settings differ). In addition the presence of constraints introduces additional difficulties in our setting. While price competition makes sense in industries where quantities produced can be easily adjusted to realized demand (such as the software industry or financial services), it may not be as realistic in industries where production capacities must be planned in advance. In these industries (such as electricity markets), quantity competition prevails (see [46] for a comparison between the two settings). Moreover, some two-stage competition games where capacity decisions are made first, followed by price competition can be recast as a one stage quantity competition[1] ([80], [52]). Kreps and Scheinkman [80] demonstrate that Bertrand competition requires a careful timing where firms decide their production quantity only after competing through prices and observing demand. For situations, where firms have to pre-commit their production capacity before competing through prices, Cournot is a better suited model.

In [60], Guo and Yang study the loss of social surplus for quantity competition. They consider unconstrained firms (firms face no production constraints) selling a homogenous good and their bound is in terms of market shares at the oligopoly equilibrium. In contrast, firms in our model sell differentiated products and face production constraints. Moreover, our surplus and profit bounds only depend on the number of firms, the number of products sold and the intensity of competition (defined in Section 2.2) which are intrinsic characteristics of the market.

The quantity competition model adopted in this chapter is drawn from [135] (Chap. 4). Our goal is to understand how competition affects social surplus and firms' profit. We adopt affine demand and cost functions. The affine demand model is widely employed in the economics literature to represent oligopoly markets ([119], [117], [120], [40], [125], [13], [48]). Yet, most of our results extend to non-linear demand as well and we point out these extensions where appropriate. Since most insights are the same in the linear and non-linear case, the chapter focuses on the linear setting to keep proofs simple.

[1] Whether or not such a two-stage game can be recast as a one-stage quantity competition depends on the rationing rule used to allocate spill-over demand of one firm which has reached capacity to the competitors.

### 2.1.2 Contributions

We consider a quantity (Cournot) competition model where firms sell differentiated substitute products (as opposed to a homogeneous good). This means that customers may have a preference for certain products, that they may be more price sensitive for some products than for others, and that the degree of substitution between products need not be identical for different pair of products. Each firm can sell multiple products.‘

Our model includes a variety of constraints, decided or faced by the firms, on their levels of production. In the context of oligopolistic competition, this work is, to the best of our knowledge, the first attempt to analyze the effect of production constraints.

As an example of the practical relevance of our setting, take the US wireless providers market. The four biggest carriers, namely Verizon, AT&T, Sprint and T-Mobile control together 89% of the cellphone market [133]. They provide substitutable differentiated products as customers may prefer one carrier over the others. Each carrier offers multiple products in the form of different phone plans: a customer must choose one of the phone plans and he might decide on the phone plan depending on prices. Each carrier faces service constraints as the wireless network can only handle so many users, talking and exchanging data, at a time. And, as carriers' advertisements often show, they compete for market share (competition in quantity) even more than they compete on prices. This is just one example of industry where the insights of this chapter could be applied towards better regulations.

We quantify both the loss of surplus and loss of profit induced by competition. We provide guarantees for the worst-case losses as a function of the number of firms, the number of products and competition intensity. In particular, the chapter establishes that none of these parameters influence the worst-case loss of surplus. In contrast, the loss of firms' total profit increases as more firms or more products are introduced in the market. Similarly, the more intense competition is, the more profit firms lose. If instead of independent firms, this model represents independent branches of a larger firm, the loss of total profit indicates how much the parent firm loses by not coordinating its branches. In that context, the loss of profit quantifies cannibalization.

**Structure of the chapter**

In Section 2.2, we describe the model and assumptions we impose. We then analyze the loss of surplus and the loss of profit for the case of single-product firms facing capacity constraints (see Section 2.3). We finally extend our model to study general convex production sets (see Section 2.4).

## 2.2 Model and assumptions

We consider a market of $n$ firms competing through quantities $\mathbf{d} = (\mathbf{d}_1, \cdots, \mathbf{d}_n)$. We denote the quantities produced by firm $i = 1, \ldots, n$ by $\mathbf{d}_i$. Each firm is assumed to sell multiple products, namely $m$ products and hence for each firm $\mathbf{d}_i = (d_{i1}, \cdots, d_{im})$.

As is traditional in the literature (see [135] Chap. 6), we model customers of this market via a representative consumer. This consumer values the possession of quantities $\mathbf{d}$ of products according to the quadratic utility function:

$$U(\mathbf{d}) = (\widetilde{\mathbf{p}})^T \ \mathbf{d} - 1/2 \ \mathbf{d}^T \mathbf{B} \ \mathbf{d}$$

For now, $\widetilde{\mathbf{p}}$ and $\mathbf{B}$ are just parameters of the utility function.

The consumer has to balance the utility he gets from owning these products against the money he spends to buy them. Therefore, for a fixed price vector $\mathbf{p}$, we define the consumer surplus as:

$$CS(\mathbf{d}) = U(\mathbf{d}) - \mathbf{p} \cdot \mathbf{d}$$

To decide the quantities to buy, the consumer maximizes his surplus: $\max_{\mathbf{d}} \ CS(\mathbf{d})$. Hence for a fixed $\mathbf{p}$, this consumer buys quantities $\mathbf{d}$ of products satisfying: $\mathbf{p} =$

$\widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$. This, in turn, gives rise to the affine, invertible demand function:

$$\mathbf{d}(\mathbf{p}) = \widetilde{\mathbf{d}} - \mathbf{M} \cdot \mathbf{p} = \begin{pmatrix} \widetilde{d}_{11} \\ \vdots \\ \widetilde{d}_{1m} \\ \vdots \\ \widetilde{d}_{nm} \end{pmatrix} - \begin{pmatrix} M_{11}^{11} & \dots & M_{11}^{1m} & \dots & M_{11}^{nm} \\ \vdots & \ddots & & & \vdots \\ M_{1m}^{11} & & M_{1m}^{1m} & & M_{1m}^{nm} \\ \vdots & & & \ddots & \vdots \\ M_{nm}^{11} & \dots & M_{nm}^{1m} & \dots & M_{nm}^{nm} \end{pmatrix} \cdot \begin{pmatrix} p_{11} \\ \vdots \\ p_{1m} \\ \vdots \\ p_{nm} \end{pmatrix}$$

where $\mathbf{M} = \mathbf{B}^{-1}$ and $\widetilde{\mathbf{d}} = \mathbf{M}\ \widetilde{\mathbf{p}}$. In the rest of the chapter, we use notations $\mathbf{M}_{kl}^{ij}$ and $\mathbf{B}_{kl}^{ij}$. Top indices refer to columns of the matrices while bottom indices refer to rows. For rows and columns, the first index ($i$ or $k$) designates a firm and the second index ($j$ or $l$) designates a product.

Anticipating the above behavior of customers, firms first decide production quantities, and then set prices to clear the market. The resulting vector of prices $\mathbf{p}_i(\mathbf{d}_i, \mathbf{d}_{-i})$, which depends on the quantities of firm $i$ and its competitors $-i$, is simply:

$$\mathbf{p}(\mathbf{d}) = \widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$$

Hence, $\widetilde{\mathbf{p}}$ can be interpreted as the prices consumers are willing to pay for their first unit of each product (assuming they own none of these products $\mathbf{d} = 0$). $\mathbf{B}$ represents the decrease in willingness to pay per unit as consumers buy more products.

We denote by $\mathbf{c}_i(\mathbf{d}_i)$ the vector of unit production costs for firm $i$ and we assume that it only depends on the firms' own production. The profit of a firm $\Pi_i$ is the difference between sales revenue and productions costs: $\Pi_i(\mathbf{d}_i, \mathbf{d}_{-i}) = \mathbf{d}_i \left[\mathbf{p}_i(\mathbf{d}_i, \mathbf{d}_{-i}) - \mathbf{c}_i(\mathbf{d}_i)\right]$. We define total profit as the sum of the firms' profit: $\Pi(\mathbf{d}) = \sum_{i=1}^{n} \Pi_i(\mathbf{d})$.

Another quantity of interest is total surplus which measures the benefit of a market for society as a whole. Total surplus is defined by aggregating consumer surplus and firms' profit:

$$\begin{aligned} TS(\mathbf{d}) &= CS(\mathbf{d}) + \Pi(\mathbf{d}) = U(\mathbf{d}) - \mathbf{p}(\mathbf{d}) \cdot \mathbf{d} + [\mathbf{p}(\mathbf{d}) - \mathbf{c}(\mathbf{d})] \cdot \mathbf{d} \\ &= [\widetilde{\mathbf{p}} - \mathbf{c}(\mathbf{d})]^T \cdot \mathbf{d} - 1/2 \ \mathbf{d}^T \mathbf{B} \ \mathbf{d} \end{aligned}$$

Throughout the chapter we will compare three market setups. In the Oligopoly Problem (OP), each firm maximizes its own profit selfishly by deciding its production quantities. Firm $i$ may face production constraints and we denote by $K_i$ its feasible production set. The result of this game is a Nash equilibrium where no firm has an incentive to change its production quantities assuming other firms will not deviate. Each firm $i$ maximizes its profit given the competitors' equilibrium quantities $\mathbf{d}_{-i}^{OP}$ by solving:

$$\begin{aligned} OP: \quad \mathbf{d}_i^{OP} = \underset{\mathbf{d}_i}{\operatorname{argmax}} \quad & \Pi_i(\mathbf{d}_i, \mathbf{d}_{-i}^{OP}) \\ \text{s.t.} \quad & \mathbf{d}_i \in K_i \end{aligned}$$

We compare this profit to the profit firms could get by colluding. Acting as a monopoly, these firms would optimize their total profit, giving rise to the Monopoly Problem (MP):

$$\begin{aligned} MP: \quad \mathbf{d}^{MP} = \underset{\mathbf{d}}{\operatorname{argmax}} \quad & \Pi(\mathbf{d}) \\ \text{s.t.} \quad & \mathbf{d} \in K. \end{aligned}$$

Set $K$ is simply the product space of $K_i$'s.

Finally, when looking at total surplus, we will benchmark these two setups against the fully coordinated situation. A central planner regulating production and consumption to maximize the surplus of society solves the Society Problem (SP):

$$\begin{aligned} SP: \quad \mathbf{d}^{SP} = \underset{\mathbf{d}}{\operatorname{argmax}} \quad & TS(\mathbf{d}) \\ \text{s.t.} \quad & \mathbf{d} \in K. \end{aligned}$$

The goal of this chapter is to compare the three setups above in terms of total surplus and firms' profit. We will often refer to these five quantities of interest: $\Pi(OP) = \Pi(\mathbf{d}^{OP})$, $\Pi(MP) = \Pi(\mathbf{d}^{MP})$, $TS(OP) = TS(\mathbf{d}^{OP})$, $TS(MP) = TS(\mathbf{d}^{MP})$

and $TS(SP) = TS(\mathbf{d}^{SP})$.

*Assumption* 1. The strategy space of seller $i$, $K_i$ is a convex, compact set such that $0 \in K_i$, and $K_i \subseteq \mathbb{R}_+^m$. This assumption allows for example constraints of the form $0 \leq \mathbf{d}_i \leq \mathbf{C}_i$, for all $i$, or in the multi-product case (each seller sells several products $j$) $\sum_j d_{ij} \leq C_i$.

*Assumption* 2. The demand function is affine: $\mathbf{d}(\mathbf{p}) = \widetilde{\mathbf{d}} - \mathbf{M} \cdot \mathbf{p}$, with $\mathbf{M}$ positive definite, $\text{diag}(\mathbf{M}) > 0$, and $\text{offdiag}(\mathbf{M}) \leq 0$.

We will assume, as is typical in the literature, that $U(\mathbf{d})$ is a strictly concave function, which translates into its Hessian matrix $\mathbf{B}$ being positive definite. This implies that $\text{diag}(\mathbf{M}) > 0$, that is, demand $d_{ij}(\mathbf{p})$ is a strictly decreasing function of $p_{ij}$.

This chapter focuses on markets of **gross substitute** products. The products are not identical, as customers may prefer to buy a certain product or to buy a given product from a specific firm. But they are gross substitutes to each other because customers will buy at most one of these products. Their decision for which product to buy will depend both on their initial preference and on the prices of the different products. For example, one can think of a market of gross substitute products as the market of midsize family cars. A given customer will most probably buy only one car and although he may have an initial preference for which car to get, he is likely to study many models and prices and to make a decision based on all the parameters at the end. In this case, we will therefore assume that $\text{offdiag}(\mathbf{M}) \leq 0$, i.e., if a firm increases its price for a certain product, some of its demand transfers to other products sold by the same firm and some of it transfers to competitors.

*Assumption* 3. $\mathbf{M}$ is a symmetric, diagonally-dominant M-matrix.

Symmetry is a consequence of the representative consumer utility assumption. Indeed, in this framework, $\nabla_{\mathbf{d}}\, \mathbf{p}(\mathbf{d}) = -\mathbf{B} = \mathrm{H}_U$, the Hessian of utility function $U$, which is a symmetric matrix.

$\mathbf{M}$ is strictly diagonally-dominant means: $M_{ij}^{ij} > \sum_{kl \neq ij} |M_{kl}^{ij}|$. This assumption is applicable to markets where total demand is decreasing with prices. More explicitly, consider the change in total market demand as a result of a unit increase in the price charged by firm i for product j, holding all other prices constant. This change is equal to $-M_{ij}^{ij} + \sum_{kl \neq ij} |M_{kl}^{ij}|$. Column diagonal dominance indicates a negative change. Under these assumptions, matrix $\mathbf{M}$ belongs to the class of M-matrices, also referred to as Stieltjes matrices, which have several useful properties and are in particular positive definite and hence invertible. We refer the reader to Horn and Johnson [68] for definition and properties of M-matrices.

Market Power:

We define market power $r_{ij}$ of firm $i$ for product $j$, and the maximum market power $r$ as:

$$r_{ij} = \sum_{kl \neq ij} |M_{kl}^{ij}|/M_{ij}^{ij} \quad \text{and} \quad r = \max_{ij} r_{ij}$$

Assumption 13, guarantees that $r \in [0, 1)$. Market power measures the intensity of competition in a market. The concept was first introduced in the context of oligopoly pricing by Farahat and Perakis (see [46]) for a market with price competition and gross substitute products. $r$ represents the fraction of customers lost by a specific firm for a given product (following an increase of its price) that will stay in the market (by buying another product from the same firm or a product from another firm). When $r$ is close to 0, decisions of a firm have very little influence on the demand seen by other firms. In the extreme case $r = 0$, matrix $\mathbf{M}$ is diagonal and the market behaves like several $(nm)$ separate monopolies. In this case, the objectives of the oligopoly and the monopoly are identical and therefore there is no loss of profit due to competition. On the other hand, when $r_{ij} = 1$, $\forall ij$ the total demand in the market remains constant, independent of the prices charged by the firms. The oligopoly problem in this setting will probably not be very efficient since it does not account at all for externalities.

*Assumption* 4. We restrict attention to constant per unit production costs $\mathbf{c}$.

Modulo some small technical assumptions, the results of the chapter still hold true for linear per unit costs but it makes notations more complex. We simplify notations by combining unit prices and unit production costs into a unit profit. We denote by $\widetilde{\mathbf{p}}$ the constant term of the price function $\mathbf{p}(\mathbf{d}) = \widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$ and by $\bar{\mathbf{p}} = \widetilde{\mathbf{p}} - \mathbf{c}$ the constant term of the per-unit profit function $\pi(\mathbf{d}) = \bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$. Similarly, we introduce $\widetilde{\mathbf{d}} = \mathbf{M} \cdot \widetilde{\mathbf{p}}$ and $\bar{\mathbf{d}} = \mathbf{M} \cdot \bar{\mathbf{p}}$.

*Assumption* 5. We assume $\bar{\mathbf{d}} = \mathbf{d}(\mathbf{c}) > 0$.

This assumption means that every firm can make a non-negative profit when all the firms charge a "sufficiently small" price. When all the firms set their unit profit to zero (price at $\mathbf{c}$), every firm observes some positive demand $\bar{\mathbf{d}}$. This, in turn, implies that the per-unit profit for a demand of zero is positive for all the firms: $\bar{\mathbf{p}} = \mathbf{B} \cdot \bar{\mathbf{d}} > 0$ since $\mathbf{B}$ is the inverse of an M-matrix ($\mathbf{M}$) and is hence componentwise non-negative.

Non-linear demand: We summarize below the set of assumptions needed to extend our results to the non-linear case.

*Assumption* 6. We assume $\mathbf{d}(0) > 0$ and $\mathbf{d}(\mathbf{c}) > 0$.

*Assumption* 7. Each demand function $d_i$ is a continuous, twice differentiable function with respect to the prices $\mathbf{p}$ and $\frac{\partial d_i}{\partial p_i} < 0$, $\frac{\partial d_i}{\partial p_j} > 0$ for $j \neq i$. The products are hence gross-substitutes.

*Assumption* 8. Demand arises from a representative consumer utility framework with concave utility $U(\mathbf{d})$. The negative of the Jacobian matrix of the demand $-\mathbf{Jd}(\mathbf{p})$ is a symmetric and strictly diagonally dominant matrix for all price vectors $\mathbf{p}$.

*Assumption* 9. The price function $\mathbf{p}_i$ is a concave function of the demand vector $\mathbf{d}$.

As a result of the previous assumptions, the negative of the Jacobian matrix of the demand $-\mathbf{Jd}(\mathbf{p})$ is positive definite for all $\mathbf{p}$. It hence satisfies the matrix similarity property (see [123] for more information). There exist a constant $A \geq 1$ such that for all $\mathbf{w}$ and $\mathbf{p}$, $\mathbf{p}'$:

$$A\ \mathbf{w}^T\ (-\mathbf{Jd}(\mathbf{p}))\ \mathbf{w} \geq \mathbf{w}^T\ (-\mathbf{Jd}(\mathbf{p}'))\ \mathbf{w} \geq \frac{1}{A}\ \mathbf{w}^T\ (-\mathbf{Jd}(\mathbf{p}))\ \mathbf{w}$$

The same property holds for the inverse $-\mathbf{J}P(\mathbf{d})$.

Notation:

- $\mathbf{B}^{\mathbf{Bdiag}}$ is the block diagonal matrix of matrix $\mathbf{B}$, each block corresponding to a firm $i$.
- $\mathbf{\Gamma}$ is the diagonal matrix made of the diagonal of $\mathbf{B}$ (the equivalent of $\mathbf{B}^{\mathbf{Bdiag}}$ for a single product per firm).
- $\mathbf{B}^{\mathbf{BOffdiag}}$ is the off block diagonal matrix: $\mathbf{B} - \mathbf{B}^{\mathbf{Bdiag}}$.
- $\mathbf{M}_{kl}^{ij}$ and $\mathbf{B}_{kl}^{ij}$. Top indices refer to columns. Bottom indices refer to rows. First index ($i$ or $k$) designates a firm. Second index ($j$ or $l$) designates a product.

Existence and Uniqueness: Since $\mathbf{B}$ is positive definite, the Monopoly and Society problems are simple optimization problems with quadratic strictly-concave objective functions. Since constraint set $K$ is a product space of $K_i$'s (each firm faces an independent constraint set), the Oligopoly problem can also be reformulated as an optimization problem on $K$. It can be checked that this new optimization problem also has a quadratic strictly-concave objective function. Constraint set $K$ being compact and convex, we are guaranteed existence and uniqueness of $\mathbf{d}^{OP}$, $\mathbf{d}^{MP}$, and $\mathbf{d}^{SP}$. Conditions for existence and uniqueness of an equilibrium such as the one resulting from the Oligopoly problem are studied by Rosen in [113].

## 2.3 Loss of total surplus and profit under production capacities

In this section, we specialize the constraints to model production capacities. Hence, sets $K_i$ take the form $0 \leq \mathbf{d}_i \leq \mathbf{C}_i \leq \bar{\mathbf{d}}_i$ where $\mathbf{C}_i$ represent the capacities. The Lagrange multipliers associated with these upper capacity constraints will be denoted by $\boldsymbol{\lambda}^{OP}$, $\boldsymbol{\lambda}^{MP}$ and $\boldsymbol{\lambda}^{SP}$ for the oligopoly, monopoly and society problems respectively. We also restrict the analysis to the single-product per firm case. We will mention conditions under which our results extend to multi-product firms.

### 2.3.1 Loss of total surplus

We study first the loss of total surplus for the entire society composed by the firms and the consumers. We remind here the expression of the total surplus:

$$TS(\mathbf{d}) = (\bar{\mathbf{p}})^T\ \mathbf{d} - 1/2\ \mathbf{d}^T\mathbf{B}\ \mathbf{d}$$

We would like to compare the surplus of society when firms compete against each other (oligopoly problem) versus when they are colluding (monopoly problem) but not coordinating with consumers. We benchmark these two cases against the maximum attainable surplus of society reached when both, firms and consumers are controlled by a central planner (society problem).

Since $SP$ is, by definition, the optimal solution for the total surplus problem, we of course have that the ratios $TS(OP)/TS(SP)$, $TS(MP)/TS(SP)$ are less than or equal to 1. We wish to establish a lower bound for these ratios that will allow us to see how $TS(OP)$ and $TS(MP)$ relate to $TS(SP)$ as well as with each other.

**Lemma 2.1.** *In a market with differentiated substitute products, a single product per firm and separate capacity constraints for each product, colluding firms always sell less quantity of each product than if they compete freely:* $\mathbf{d}^{MP} \leq \mathbf{d}^{OP}$.

The proof is left in appendix A.1.

**Theorem 2.1.** *For differentiated substitute products, with a single product per firm and separate capacity constraints for each product, free competition between firms is always better for consumers (in terms of consumer surplus) than allowing the firms to collude: $CS(OP) \geq CS(MP)$.*

*Proof.* The consumer surplus can simply be written as:

$$\begin{aligned} CS(\mathbf{d}) &= U(\mathbf{d}) - \mathbf{p}(\mathbf{d})\mathbf{d} \\ &= \tilde{\mathbf{p}}\mathbf{d} - \frac{1}{2}\,\mathbf{dBd} - \tilde{\mathbf{p}}\mathbf{d} + \mathbf{dBd} \\ &= \frac{1}{2}\,\mathbf{dBd} \end{aligned}$$

It follows that:

$$\begin{aligned} CS(OP) - CS(MP) &= \frac{1}{2}\,\mathbf{d}^{OP}\mathbf{B}\mathbf{d}^{OP} - \frac{1}{2}\,\mathbf{d}^{MP}\mathbf{B}\mathbf{d}^{MP} \\ &= \frac{1}{2}\,\underbrace{(\mathbf{d}^{OP} + \mathbf{d}^{MP})}_{\geq 0}\,\underbrace{\mathbf{B}}_{\geq 0}\,\underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{MP})}_{\geq 0} \\ CS(OP) - CS(MP) &\geq 0 \end{aligned}$$

□

**Theorem 2.2.** *For differentiated substitute products, with a single product per firm and separate capacity constraints for each product, free competition between firms always achieves greater total surplus for society than allowing them to collude: $TS(OP) \geq TS(MP)$.*

**Theorem 2.3.** *In a market with differentiated substitute products, a single product per firm and separate capacity constraints for each product, allowing firms to collude reduces total surplus for society by at most 25%: $TS(MP) \geq 3/4\ TS(SP)$.*

Tightness in Theorem 2.1 and 2.2 is trivially achieved when matrix $\mathbf{M}$ is diagonal. In this case, firms do not impact each other and oligopoly and monopoly solutions are the same.

In Theorem 2.3, tightness is achieved by removing the capacity constraints. Without constraints $\mathbf{d}^{MP} = 1/2\ \bar{\mathbf{d}}$ and $\mathbf{d}^{SP} = \bar{\mathbf{d}}$, leading to $TS(MP)/TS(SP) = 3/4$.

It is interesting to note that the bounds obtained in Theorem 2.1, 2.2 and 2.3, are the same as in the unconstrained case. This means that adding simple capacity constraints to the production set of each firm does not worsen efficiency. This is to be contrasted with the results of the next section where adding more general convex constraints lowers the bounds.

*Proof.* (Theorem 2.2)

$$\begin{aligned}
TS(OP) - TS(MP) &= \bar{\mathbf{p}}\ \mathbf{d}^{OP} - \frac{1}{2}\mathbf{d}^{OP}\mathbf{B}\mathbf{d}^{OP} - \bar{\mathbf{p}}\ \mathbf{d}^{MP} + \frac{1}{2}\mathbf{d}^{MP}\mathbf{B}\mathbf{d}^{MP} \\
&= \bar{\mathbf{d}}\ \mathbf{B}\ (\mathbf{d}^{OP} - \mathbf{d}^{MP}) - \frac{1}{2}(\mathbf{d}^{OP} + \mathbf{d}^{MP})\ \mathbf{B}\ (\mathbf{d}^{OP} - \mathbf{d}^{MP}) \\
&= \underbrace{\left(\bar{\mathbf{d}} - \frac{\mathbf{d}^{OP} + \mathbf{d}^{MP}}{2}\right)}_{\geq 0}\ \underbrace{\mathbf{B}}_{\geq 0}\ \underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{MP})}_{\geq 0} \\
TS(OP) - TS(MP) &\geq 0
\end{aligned}$$

The last inequality comes from the result $\mathbf{d}^{MP} \leq \mathbf{d}^{OP}$ that we just established above. □

*Proof.* (Theorem 2.3)

First let's formulate the society problem (SP) under capacity constraints. It can be written as:

$$\begin{aligned}
\max_{\mathbf{d}}\quad & \bar{\mathbf{p}}\ \mathbf{d} - \tfrac{1}{2}\ \mathbf{d}\ \mathbf{B}\ \mathbf{d} \\
\text{s.t.}\quad & 0 \leq \mathbf{d} \leq \mathbf{C} \leq \bar{\mathbf{d}}
\end{aligned}$$

The corresponding KKT conditions for problem (SP) are:

$$\bar{\mathbf{p}} - \mathbf{B}\ \mathbf{d}^{SP} - \boldsymbol{\lambda}^{SP} + \boldsymbol{\mu}^{SP} = 0 \qquad \left\{\begin{array}{c} (\boldsymbol{\lambda}^{SP})^T(\mathbf{C} - \mathbf{d}^{SP}) = 0 \\ \boldsymbol{\lambda}^{SP} \geq 0 \\ \mathbf{d}^{SP} \leq \mathbf{C} \leq \bar{\mathbf{d}} \end{array}\right. \qquad \left\{\begin{array}{c} (\boldsymbol{\mu}^{SP})^T\mathbf{d}^{SP} = 0 \\ \boldsymbol{\mu}^{SP} \geq 0 \\ \mathbf{d}^{SP} \geq 0 \end{array}\right.$$

Step 1: $\boldsymbol{\mu}^{SP} = 0$ and $\mathbf{d}^{SP} = \mathbf{C}$

The proof is left in appendix A.2.

Step 2: *Constraints Splitting and Initial Calculations* ($\boldsymbol{\lambda}^{MP}$ is defined in appendix A.1)

Let $K_2 = \{\text{Set of active constraints for the monopoly problem}\} = \{i = 1, \ldots, n, \quad \lambda_i^{MP} > 0\}$. We denote by $K_2^c$ the complement set of $K_2$ and by $\mathbf{H}_{AB}$ and $\mathbf{u}_A$ the restrictions of matrix $\mathbf{H}$ and vector $\mathbf{u}$ to rows indexed by $A$ and columns indexed by $B$. Since $K_2$ is the set of active capacity constraints for problem (MP),

$$\mathbf{d}^{MP} = \begin{pmatrix} d_{K_2}^{MP} \\ d_{K_2^c}^{MP} \end{pmatrix} = \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix}.$$

$$\begin{aligned}
TS(MP) - \frac{3}{4}\,TS(SP) &= \bar{\mathbf{p}}\;\mathbf{d}^{MP} - \frac{1}{2}\;\mathbf{d}^{MP}\;\mathbf{B}\;\mathbf{d}^{MP} - \frac{3}{4}\;\bar{\mathbf{p}}\;\mathbf{d}^{SP} + \frac{3}{8}\;\mathbf{d}^{SP}\;\mathbf{B}\;\mathbf{d}^{SP} \\
&= \bar{\mathbf{p}} \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix} - \frac{1}{2}(c_{K_2}\; d_{K_2^c}^{MP})\;\mathbf{B} \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix} - \frac{3}{4}\bar{\mathbf{p}} \begin{pmatrix} c_{K_2} \\ c_{K_2^c} \end{pmatrix} \\
&\quad + \frac{3}{8}(c_{K_2}\; c_{K_2^c})\;\mathbf{B} \begin{pmatrix} c_{K_2} \\ c_{K_2^c} \end{pmatrix} \\
&\geq \bar{\mathbf{p}} \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix} - \frac{1}{2}(c_{K_2}\; d_{K_2^c}^{MP})\;\mathbf{B} \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix} - \frac{3}{4}\bar{\mathbf{p}} \begin{pmatrix} c_{K_2} \\ \bar{d}_{K_2^c} \end{pmatrix} \\
&\quad + \frac{3}{8}(c_{K_2}\; \bar{d}_{K_2^c})\;\mathbf{B} \begin{pmatrix} c_{K_2} \\ \bar{d}_{K_2^c} \end{pmatrix}
\end{aligned}$$

The last inequality comes from the fact that $c_{K_2^c} \leq \bar{d}_{K_2^c}$ and that:

$$\nabla[\bar{\mathbf{p}}\;\mathbf{d} - \frac{1}{2}\mathbf{dBd}] = \bar{\mathbf{p}} - \mathbf{Bd} = \mathbf{B}(\bar{\mathbf{d}} - \mathbf{d}) \geq 0$$

Back to our calculations:

$$
\begin{aligned}
TS(MP) - \frac{3}{4}\,TS(SP) \;\; \geq \;\; & \bar{d}_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} + \bar{d}_{K_2}\;\mathbf{B}_{K_2K_2^c}\;d_{K_2^c}^{MP} + \bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2}\;c_{K_2} \\
+ \;\; & \bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2^c}\;d_{K_2^c}^{MP} - \frac{1}{2}c_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} - c_{K_2}\;\mathbf{B}_{K_2K_2^c}\;d_{K_2^c}^{MP} \\
- \;\; & \frac{1}{2}d_{K_2^c}^{MP}\;\mathbf{B}_{K_2^cK_2^c}\;d_{K_2^c}^{MP} - \frac{3}{4}\bar{d}_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} - \frac{3}{4}\bar{d}_{K_2}\;\mathbf{B}_{K_2K_2^c}\;\bar{d}_{K_2^c} \\
- \;\; & \frac{3}{4}\bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2}\;c_{K_2} - \frac{3}{4}\bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2^c}\;\bar{d}_{K_2^c} + \frac{3}{8}c_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} \\
+ \;\; & \frac{3}{4}c_{K_2}\;\mathbf{B}_{K_2K_2^c}\;\bar{d}_{K_2^c} + \frac{3}{8}\bar{d}_{K_2^c}\;\mathbf{B}_{K_2K_2^c}\;\bar{d}_{K_2^c}
\end{aligned}
$$

Grouping terms we get:

$$
\begin{aligned}
TS(MP) - \frac{3}{4}\,TS(SP) \;\; \geq \;\; & \frac{1}{4}\bar{d}_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} - \frac{1}{8}c_{K_2}\;\mathbf{B}_{K_2K_2}\;c_{K_2} + \bar{d}_{K_2}\;\mathbf{B}_{K_2K_2^c}\;d_{K_2^c}^{MP} \\
& -c_{K_2}\;\mathbf{B}_{K_2K_2^c}\;d_{K_2^c}^{MP} + c_{K_2}\;\mathbf{B}_{K_2K_2^c}\;\bar{d}_{K_2^c} - \frac{3}{4}\bar{d}_{K_2}\;\mathbf{B}_{K_2K_2^c}\;\bar{d}_{K_2^c} \\
& +\bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2^c}\;d_{K_2^c}^{MP} - \frac{1}{2}d_{K_2^c}^{MP}\;\mathbf{B}_{K_2^cK_2^c}\;d_{K_2^c}^{MP} - \frac{3}{8}\bar{d}_{K_2^c}\;\mathbf{B}_{K_2^cK_2^c}\;\bar{d}_{K_2^c}
\end{aligned}
$$

Step 3: *Using monopoly (MP) KKT conditions in inequality*

As shown in appendix A.1, the monopoly KKT conditions imply that:

$$\bar{\mathbf{p}} - 2\;\mathbf{B}\;\mathbf{d}^{MP} - \boldsymbol{\lambda}^{MP} = 0$$

Restricting attention to the set $K_2^c$ of inactive constraints ($\lambda_{K_2^c}^{MP} = 0$):

$$
\begin{aligned}
& \bar{\mathbf{p}}_{K_2^c} - 2\;\mathbf{B}_{K_2^c}\begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix} = 0 \\
\Rightarrow \quad & d_{K_2^c}^{MP} = (\mathbf{B}_{K_2^cK_2^c})^{-1}\left(\frac{1}{2}\;\bar{\mathbf{p}}_{K_2^c} - \mathbf{B}_{K_2^cK_2}\;c_{K_2}\right) \\
\Rightarrow \quad & d_{K_2^c}^{MP} = \frac{1}{2}\;\bar{\mathbf{d}}_{K_2^c} + (\mathbf{B}_{K_2^cK_2^c})^{-1}\;\mathbf{B}_{K_2^cK_2}\;\left(\frac{1}{2}\;\bar{\mathbf{d}}_{K_2} - c_{K_2}\right) \qquad (2.1)
\end{aligned}
$$

Now, replacing the expression of $d^{MP}_{K_2^c}$ in the rest of the KKT conditions (set $K_2$):

$$\begin{aligned} & \bar{\mathbf{p}}_{K_2} - 2\,\mathbf{B}_{K_2} \begin{pmatrix} c_{K_2} \\ d^{MP}_{K_2^c} \end{pmatrix} \geq 0 \\ \Rightarrow \quad & [\mathbf{B}_{K_2K_2} - \mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}]\,(\bar{d}_{K_2} - 2c_{K_2}) \geq 0 \end{aligned} \tag{2.2}$$

Using expression (2.1), we simplify the relation between $TS(MP)$ and $TS(SP)$:

$$\begin{aligned} TS(MP) - \frac{3}{4}\,TS(SP) \;\geq\; & \frac{3}{8}\,\bar{d}_{K_2}\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,\bar{d}_{K_2} \\ + & \frac{1}{4}\,\bar{d}_{K_2}\,\mathbf{B}_{K_2K_2}\,c_{K_2} - \bar{d}_{K_2}\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,c_{K_2} \\ - & \frac{1}{8}\,c_{K_2}\,\mathbf{B}_{K_2K_2}\,c_{K_2} + \frac{1}{2}\,c_{K_2}\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,c_{K_2} \end{aligned}$$

Regrouping terms to make expression (2.2) appear:

$$\begin{aligned} TS(MP) - \frac{3}{4}\,TS(SP) \;\geq\; & \frac{3}{8}\,\bar{d}_{K_2}\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,\bar{d}_{K_2} \\ - & \frac{3}{4}\,\bar{d}_{K_2}\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,c_{K_2} + \frac{3}{8}\,c_{K_2}\,\mathbf{B}_{K_2K_2}\,c_{K_2} \\ + & \frac{1}{4}\,c_{K_2}\,\underbrace{[\mathbf{B}_{K_2K_2} - \mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}]\,(\bar{d}_{K_2} - 2c_{K_2})}_{\geq 0} \end{aligned}$$

Finally:

$$\begin{aligned} TS(MP) - \frac{3}{4}\,TS(SP) \;\geq\; & \frac{3}{8}\,(\bar{d}_{K_2} - c_{K_2})\,\mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}\,(\bar{d}_{K_2} - c_{K_2}) \quad \geq 0 \\ + & \frac{3}{8}\,c_{K_2}\,[\mathbf{B}_{K_2K_2} - \mathbf{B}_{K_2K_2^c}\,(\mathbf{B}_{K_2^cK_2^c})^{-1}\,\mathbf{B}_{K_2^cK_2}]\;c_{K_2} \quad \geq 0 \end{aligned}$$

The top term is non-negative simply because $B_{K_2^cK_2^c}$ is still an inverse M-matrix (see [72]) and as such its inverse is positive semi-definite. The second term is non-negative because:

$$\mathbf{B}_{K_2K_2} - \mathbf{B}_{K_2K_2^c}\left(\mathbf{B}_{K_2^cK_2^c}\right)^{-1}\mathbf{B}_{K_2^cK_2} = \left(M_{K_2K_2}\right)^{-1}$$

is also positive semi-definite.

Hence, it follows that: $TS(MP) \geq \frac{3}{4}\, TS(SP)$. $\square$

Extension 1: Multi-product firms

We restricted our analysis above to markets with a single product per firm. In fact, our analysis extends mutatis mutandis to markets with multi-product firms as long as $\mathbf{B} + \mathbf{B}^{\mathbf{Bdiag}}$ stays an inverse M-matrix. Starting with $\mathbf{M}$ being an M-matrix it is not always true that $\mathbf{B} + \mathbf{B}^{\mathbf{Bdiag}}$ will be an inverse M-matrix.

Counterexample: We take 2 firms selling 2 products. Matrix $\mathbf{B}$ below is an inverse M but $\mathbf{B}+\mathbf{B}^{\mathbf{Bdiag}}$ is not an inverse M-matrix (here $\mathbf{B}^{\mathbf{Bdiag}}$ is the block-diagonal matrix with 2 by 2 diagonal blocks).

$$\mathbf{B} = \begin{pmatrix} 40 & 3 & 18 & 8 \\ 3 & 9 & 1.75 & 2 \\ 18 & 1.75 & 52 & 9 \\ 8 & 2 & 9 & 23 \end{pmatrix}$$

But under some additional assumptions, it is possible to guarantee this property. In particular, if $\mathbf{B}$ is a semi-uniform (off-diagonal coefficients all the same) inverse M-matrix, using properties of strict ultrametricity (see [110]) it can be shown that $\mathbf{B} + \mathbf{B}^{\mathbf{Bdiag}}$ will be an inverse M-matrix. In that case, all our results hold.

Extension 2: General polyhedral constraints

We only analyzed here the effect of separate upper capacity constraints on each product. We believe our proofs can be easily extended to include separate minimum production constraints for each product. All our results should still hold true. Moreover, extensive random numerical simulations suggest that our results are still true under general polyhedral set.

In conclusion, this implies that the traditional economic result that competition benefits society holds true no matter how many firms compete and under very general constraint sets.

### 2.3.2 Loss of firms profit

Our goal in this subsection is to study the drivers of the loss of total profit in the market due to competition between firms. Hence, we use as a benchmark the total profit in the market when firms are colluding. That is, we compare total profit under the oligopoly solution versus the monopoly solution, i.e., $\Pi(OP)/\Pi(MP) = \sum_i \Pi_i(\mathbf{d}^{OP})/\sum_i \Pi_i(\mathbf{d}^{MP}) \leq 1$. This section establishes a lower bound for this ratio.

We first consider a uniform market with identical firms. $\mathbf{M}$ is assumed to be a uniform matrix: its diagonal coefficients are all equal to 1 (after normalization) and its off-diagonal coefficients are all equal (to $-\alpha$). Firms differ only through their initial demand $\bar{d}_i$. We also get rid of production constraints as we believe capacity constraints can only increase the ratio $\Pi(OP)/\Pi(MP)$. After studying the uniform unconstrained case, we will show numerically that the bound extends to the general case.

**Theorem 2.4.** *In a uniform market with differentiated substitute products, a single product per firm and no production constraints, forcing firms to compete reduces their profit (compared to the collusion case) by:*

$$\Pi(OP)/\Pi(MP) \geq \min \left\{ \begin{array}{c} \frac{4(n-1+r)((n-1)(1-r)+r)}{((n-1)(2-r)+2r)^2} \\ \frac{4}{n}\left[\left(\frac{n-1+r}{(n-1)(2-r)+2r}\right)^2 + (n-1)\left(\frac{(n-1)(1-r)}{2(n-1)(1-r)+r}\right)^2\right] \end{array} \right\}$$

.

We would like to emphasize here that this bound which is tight (as shown after the proof) gives us an exact characterization of the maximum loss of profit as a function of the number of firms $n$ in the market and the intensity of competition (measured by the market power $r$). This is significant as both $n$ and $r$ can be "easily" determined from the market characteristics. If instead of independent firms, we think of each firm above as a branch of a decentralized larger firm taking independent decisions, the ratio above measures how much profit the firm is losing as a result of not coordinating its branches. In some sense, it is a measure of cannibalization.

*Proof.* Without production constraints, the KKT conditions of Appendix A.1 simplify to $\mathbf{d}^{OP} = (\mathbf{B}+\mathbf{\Gamma})^{-1}\bar{\mathbf{p}}$ and $\mathbf{d}^{MP} = 1/2\ \mathbf{M}\ \bar{\mathbf{p}}$. Substituting $\mathbf{B}\bar{\mathbf{d}}$ for $\bar{\mathbf{p}}$ and computing $\Pi(\mathbf{d}) = \mathbf{d}\cdot\{\bar{\mathbf{p}} - \mathbf{B}\cdot\mathbf{d}\}$, we get:

$$\frac{\Pi(OP)}{\Pi(MP)} = \frac{4\ \bar{\mathbf{d}}^T(I+\mathbf{\Gamma}\mathbf{M})^{-1}\ \mathbf{\Gamma}\ (I+\mathbf{M}\mathbf{\Gamma})^{-1}\bar{\mathbf{d}}}{\bar{\mathbf{d}}\ \mathbf{B}\ \bar{\mathbf{d}}}$$

In the uniform case, skipping some calculations (provided in appendix A.3), the numerator and denominator above can be diagonalized with only two eigenvalues. Let's call $\breve{\mathbf{d}}$ the vector whose components are the eigenvectors of $\mathbf{M}$, and $[\breve{\rho}_1, \breve{\rho}_2]$ the two eigenvalues of: $(I+\mathbf{\Gamma}\mathbf{M})^{-1}\ \mathbf{\Gamma}\ (I+\mathbf{M}\mathbf{\Gamma})^{-1}$.

- $\breve{\rho}_1 = \frac{(1+\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-n\alpha)^2(1+\alpha-n\alpha)}$

- $\breve{\rho}_2 = \frac{(1+\alpha-n\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-2n\alpha)^2(1+\alpha)}$

The ratio of profits becomes:

$$\frac{\Pi(OP)}{\Pi(MP)} = \frac{4\ (\breve{\rho}_1\breve{d}_1^2 + \breve{\rho}_2\sum_{i=2}^n \breve{d}_i^2)}{\frac{1}{1+\alpha-n\alpha}\breve{d}_1^2 + \frac{1}{1+\alpha}\sum_{i=2}^n \breve{d}_i^2}$$

To normalize this ratio, let's define:

$$\breve{\mathbf{d}} = \begin{pmatrix} \frac{1}{\sqrt{1+\alpha-n\alpha}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{1+\alpha}} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\sqrt{1+\alpha}} \end{pmatrix} \breve{\mathbf{d}} = \begin{pmatrix} \frac{1}{\sqrt{1+\alpha-n\alpha}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{1+\alpha}} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\sqrt{1+\alpha}} \end{pmatrix} \Delta^T \bar{\mathbf{d}}$$

We can rewrite the profit ratio as:

$$\frac{\Pi(OP)}{\Pi(MP)} = \frac{\breve{\rho}_1 \breve{d}_1^2 + \breve{\rho}_2 \sum_{i=2}^n \breve{d}_i^2}{\sum_{i=1}^n \breve{d}_i^2} \tag{2.3}$$

where

- $\breve{\rho}_1 = \frac{4\ (1+\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-n\alpha)^2}$
- $\breve{\rho}_2 = \frac{4\ (1+\alpha-n\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-2n\alpha)^2}$

By hypothesis, we have $\alpha \geq 0$ and $(n-1)\,\alpha \leq 1$. Using these assumptions, it can then easily be shown that: $\breve{\rho}_1, \breve{\rho}_2 \geq 0$ (and that both their numerator and denominator are non-negative). Moreover, by developing the numerator and the numerator of these eigenvalues and using the assumptions above, we can also show that:

- $\breve{\rho}_1 \geq \breve{\rho}_2$ when $r = \alpha(n-1) \geq \frac{(n-1)(n-2)}{(n-1)(n-2)+1}$
- $\breve{\rho}_2 \geq \breve{\rho}_1$ otherwise.

We are looking for the worst possible $\breve{\mathbf{d}}$ of ratio (2.3), satisfying $\bar{\mathbf{d}} \geq 0$. This problem is Norm-2 invariant, so that we can restrict our search to vectors satisfying $||\breve{\mathbf{d}}||_2^2 = 1$.

We can write this problem as an optimization:

$$\min_{\breve{\mathbf{d}}} \quad \breve{\rho}_1 \breve{d}_1^2 + \breve{\rho}_2 \sum_{i=2}^{n} \breve{d}_i^2$$

$$\text{s.t.} \quad \begin{cases} \bar{\mathbf{d}} = \Delta \begin{pmatrix} \sqrt{1+\alpha-n\alpha} & 0 & \dots & 0 \\ 0 & \sqrt{1+\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{1+\alpha} \end{pmatrix} \breve{\mathbf{d}} \geq 0 \\ \\ \|\breve{\mathbf{d}}\|_2^2 = 1 \end{cases}$$

This problem can be reformulated:

$$\min_{\breve{\mathbf{d}}} \quad \breve{\rho}_1 + (\breve{\rho}_2 - \breve{\rho}_1) \sum_{i=2}^{n} \breve{d}_i^2$$

$$\text{s.t.} \quad \begin{cases} \bar{\mathbf{d}} = \Delta \begin{pmatrix} \sqrt{1+\alpha-n\alpha} & 0 & \dots & 0 \\ 0 & \sqrt{1+\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{1+\alpha} \end{pmatrix} \breve{\mathbf{d}} \geq 0 \\ \\ \|\breve{\mathbf{d}}\|_2^2 = 1 \end{cases}$$

In order to analyze this optimization problem, we need to spell out matrix $\Delta$. By construction, $\Delta$ is a unimodular matrix made of the eigenvectors of $\mathbf{M}$. It is:

$$\Delta = \begin{pmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \frac{-1}{\sqrt{2}} \\ \vdots & & \frac{-(n-2)}{\sqrt{(n-1)(n-2)}} & 0 & \vdots \\ \frac{1}{\sqrt{n}} & \frac{-(n-1)}{\sqrt{n(n-1)}} & 0 & \cdots & 0 \end{pmatrix}$$

If $\breve{\breve{\rho}}_2 \geq \breve{\breve{\rho}}_1$, the above minimum can be achieved by choosing $\{\breve{\breve{d}}_i = 0,\ i = 2, \ldots, n\}$. Since $\breve{\breve{\mathbf{d}}} = (1, 0, \ldots, 0)$ is a feasible solution (it corresponds to $\bar{\mathbf{d}} = \sqrt{1+\alpha-n\alpha}/\sqrt{n}\ \mathbf{e} \geq 0$), it is an optimal solution and the optimal value is $\breve{\breve{\rho}}_1$. In this case, $\Pi(OP)/\Pi(MP) \geq \breve{\breve{\rho}}_1$.

Otherwise, if $\breve{\breve{\rho}}_1 \geq \breve{\breve{\rho}}_2$, this problem is equivalent (for the search of optimal solutions) to:

$$\max_{\breve{\breve{\mathbf{d}}}} \ \sum_{i=2}^{n} \breve{\breve{d}}_i^2$$

$$\text{s.t.} \begin{cases} \bar{\mathbf{d}} = \Delta \begin{pmatrix} \sqrt{1+\alpha-n\alpha} & 0 & \cdots & 0 \\ 0 & \sqrt{1+\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{1+\alpha} \end{pmatrix} \breve{\breve{\mathbf{d}}} \geq 0 \\ \\ \|\breve{\breve{\mathbf{d}}}\|_2^2 = 1 \end{cases} \tag{2.4}$$

To ease notations, we re-scale this problem to:

$$\max_{\check{\mathbf{d}}} \quad \sum_{i=2}^{n} \check{d}_i^2$$

$$\text{s.t.} \quad \begin{cases} \bar{\mathbf{d}} = \Delta \check{\mathbf{d}} \geq 0 \\ \\ \sigma\ \check{d}_1^2 + \sum_{i=2}^{n} \check{d}_i^2 = 1 + \alpha \end{cases} \tag{2.5}$$

where $\sigma = \frac{1+\alpha}{1+\alpha-n\alpha}$ will be used in the remainder of the proof.

We now write the first order KKT conditions of optimization problem (2.5). These can be summarized as:

$$\begin{cases} \nabla_{\check{\mathbf{d}}} \mathcal{L}(\check{\mathbf{d}}, \lambda, \mu) = 2 \begin{pmatrix} 0 \\ \check{d}_2 \\ \vdots \\ \check{d}_n \end{pmatrix} + \Delta^T \lambda - 2\mu \begin{pmatrix} \sigma\ \check{d}_1 \\ \check{d}_2 \\ \vdots \\ \check{d}_n \end{pmatrix} = 0 \\ \\ \lambda \geq 0 \\ \\ \lambda^T \Delta\ \check{\mathbf{d}} = 0 \end{cases}$$

Multiplying the KKT equation above by $\Delta$, we get:

$$\Delta \nabla_{\check{\mathbf{d}}} \mathcal{L}(\check{\mathbf{d}}, \lambda, \mu) = 0$$

$$\lambda + 2\ \Delta \begin{pmatrix} -\mu\ \sigma\ \check{d}_1 \\ (1-\mu)\ \check{d}_2 \\ \vdots \\ (1-\mu)\ \check{d}_n \end{pmatrix} = 0$$

From which we get:

$$\lambda = -2\ (1-\mu)\ \Delta\ \breve{\mathbf{d}}\ +\ 2\ (1-\mu+\mu\ \sigma)\ \Delta \begin{pmatrix} \breve{d}_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

This combined with complementary slackness results in an explicit expression of $\lambda$:

- either $\Delta_i\ \breve{\mathbf{d}} > 0$ in which case $\lambda_i = 0$.
- or $\Delta_i\ \breve{\mathbf{d}} = 0$, in which case $\lambda_i = 2\ (1+\mu\ (\sigma-1))\ \frac{\breve{d}_1}{\sqrt{n}}$

On the other hand, multiplying the KKT equation by $\breve{\mathbf{d}}$ and using the initial constraints of problem (2.5), we get:

$$\begin{aligned} & \breve{\mathbf{d}}^T\ \nabla_{\breve{\mathbf{d}}}\mathcal{L}(\breve{\mathbf{d}},\lambda,\mu) = 0 \\ \Leftrightarrow\quad & 2\textstyle\sum_{i=2}^n \breve{d}_i^2 - 2\ \mu\ (1+\alpha) = 0 \\ \Leftrightarrow\quad & \mu = 1/(1+\alpha)\textstyle\sum_{i=2}^n \breve{d}_i^2 = \textstyle\sum_{i=2}^n \breve{d}_i^2 \end{aligned}$$

So our goal is to maximize the value of $\mu$.

Finally, let's denote by $A$ the set of active constraints of equation $\bar{\mathbf{d}} = \Delta\breve{\mathbf{d}} \geq 0$ in problem (2.5). Multiplying the KKT equation by $\mathbf{e_1} = (1, 0, \ldots, 0)$, we get:

$$\begin{aligned} & \tfrac{1}{\sqrt{n}}\textstyle\sum_{i\in A}\lambda_i = 2\ \mu\ \sigma\ \breve{d}_1 \\ \Leftrightarrow\quad & \tfrac{|A|}{n}\ (1+\mu\ (\sigma-1)) = \mu\ \sigma \\ \Leftrightarrow\quad & \mu = \textstyle\sum_{i=2}^n \breve{d}_i^2 = \frac{1}{\sigma\ (\frac{n}{|A|}-1)+1} \end{aligned}$$

Since we are maximizing $\mu$, we want the number of active constraints ($|A|$) as large as possible. Because $\Delta$ is invertible, the only vector satisfying $n$ active constraints is $\breve{\mathbf{d}} = 0$ which violates the other constraint $\sigma\ \breve{d}_1^2 + \sum_{i=2}^n \breve{d}_i^2 = 1 + \alpha$.

It is however possible to satisfy $n-1$ constraints in $A$ and the constraint $\sigma\ \breve{d}_1^2 + \sum_{i=2}^n \breve{d}_i^2 = 1 + \alpha$. One such feasible solution is $\breve{\mathbf{d}} = (u, -\sqrt{n-1}\ u, 0, \ldots, 0)$. (with $u = \sqrt{\frac{(1+\alpha-n\alpha)(1+\alpha)}{n(1+2\alpha-n\alpha)}}$)

Hence, we choose $|A| = n-1$ and $\mu = \sum_{i=2}^n \breve{\breve{d}}_i^2 = \frac{n-1}{\sigma+n-1}$. The corresponding objective value is $\breve{\breve{\rho}}_1 + (\breve{\breve{\rho}}_2 - \breve{\breve{\rho}}_1) \sum_{i=2}^n \breve{d}_i^2$.

So when $\breve{\breve{\rho}}_1 \geq \breve{\breve{\rho}}_2$, replacing for $\breve{\breve{\rho}}_1$, $\breve{\breve{\rho}}_2$, and $\sigma$, we obtain:

$$\Pi(OP)/\Pi(MP) \geq \frac{4}{n}\left[\frac{(1+\alpha)^2}{(2+3\alpha-n\alpha)^2} + (n-1)\frac{(1+\alpha-n\alpha)^2}{(2+3\alpha-2n\alpha)^2}\right]$$

Combining the two cases ($\breve{\breve{\rho}}_1 \geq \breve{\breve{\rho}}_2$ and $\breve{\breve{\rho}}_2 \geq \breve{\breve{\rho}}_1$) and replacing $(n-1)\alpha = r$, we obtain:

$$\Pi(OP)/\Pi(MP) \geq \min\left\{\begin{array}{c} \frac{4(n-1+r)((n-1)(1-r)+r)}{((n-1)(2-r)+2r)^2} \\ \frac{4}{n}\left[\left(\frac{n-1+r}{(n-1)(2-r)+2r}\right)^2 + (n-1)\left(\frac{(n-1)(1-r)}{2(n-1)(1-r)+r}\right)^2\right] \end{array}\right\}$$

The top part of the bound is active when $r \leq \frac{(n-1)(n-2)}{(n-1)(n-2)+1}$, the bottom part is active otherwise.

$\square$

Tightness of the bound is again easily shown. Following the proof above, we pick $\mathbf{M}$ uniform and $\bar{\mathbf{d}} = \mathbf{e}$ achieves the top part of the bound in Theorem 2.4 while $\bar{\mathbf{d}} = (0, \ldots, 0, 1)$ achieves the bottom part.

We now show numerically that the bound above holds true for non-uniform unconstrained markets. We simulate random non-uniform markets for different values of $r$ and $n$ and we compare with the bound. We tried to be as general as possible in our simulation method. Here are the steps we followed:

- We simulate separately for different number of firms $n = 2 \dots 10$.
- For each $n$, we vary the market power $r$ from 0 to 1 by 0.1 increments.
- For each pair ($n$,$r$) we repeat the experiment $50n^2$ times. We pick the number of simulations proportional to $n^2$ because we simulate each coefficient of matrix $\mathbf{M}$ randomly (leading to $\Omega(n^2)$ random coefficients). We consider the oligopoly vs monopoly profit ratio and we pick the lowest ratio of these $50n^2$ experiments.
- For each experiment, we simulate $\bar{\mathbf{d}}$ and $\mathbf{M}$ randomly.
  - Each coefficient of $\bar{\mathbf{d}}$ is random in [0,1].
  - Each off-diagonal coefficient of $\mathbf{M}$ is random between [-1,0]. We keep symmetry by taking the symmetric part of $\mathbf{M}$ here.
  - The first diagonal coefficient $\mathbf{M}_{11}^{11}$ is equal to $\sum_{kl \neq 11} |M_{kl}^{11}|/r$ to guarantee a market power of $r$
  - The other diagonal coefficients $\mathbf{M}_{ij}^{ij}$ are equal to the absolute value of the sum of the corresponding off-diagonal column coefficients divided by $r$ times some random number $rand_{ij}$ in [0,1]: $\sum_{kl \neq ij} |M_{kl}^{ij}|/(r * rand_{ij})$. This guarantees that $r_{ij} \leq r$.
  - $\bar{\mathbf{p}} = \mathbf{B}\bar{\mathbf{d}}$ by definition and these parameters entirely determine the unconstrained solution.

Each plot in the figure below corresponds to a value of $n = 2 \dots 10$ (increasing from left to right and top to bottom). The plots represent the ratio $\Pi(OP)/\Pi(MP)$ as a function of the market power $r$. The red curve is the theoretical bound, the blue one (with the square) corresponds to our simulation. Our bound seems to still hold in all cases. Tightness has already been shown above.

Finally, we analyze the effect of capacity constraints. We repeat the previous modus operandi to simulate the effect of constraints. For each pair ($n$, $r$), we generate

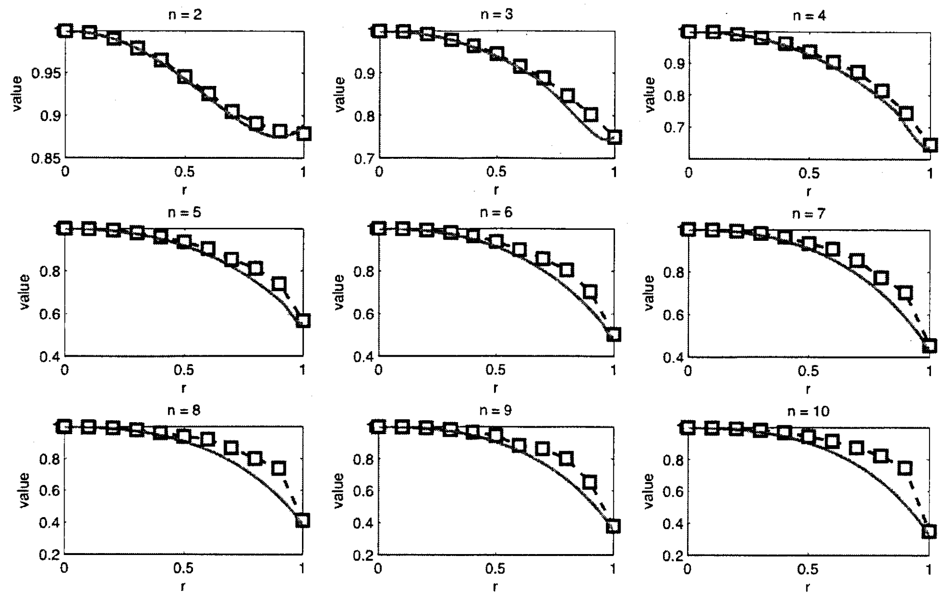


Figure 2-1: Numerical simulation of the profit ratio $\Pi(OP)/\Pi(MP)$ for general unconstrained markets (in blue) and theoretical bound for uniform markets (in red).

many experiments. We compute the profit ratio for the unconstrained case. We then randomly generate separate upper capacities $C_{ij}$ for each product (product $j$ from firm $i$) and we compute the profit ratio with capacities. The purple curve (top one) includes upper capacities, the blue one corresponds to the unconstrained case. We represent here the lowest value of the profit ratio across experiments. It appears that the worst unconstrained case is lower than the worst constrained case. Our theoretical bound should then hold true even with production capacities.

## 2.4 Loss of total surplus and profit under general constraint sets

In this section, we extend our analysis to multi-product firms facing general production constraints. Each firm sells $m$ products. $K_i$ only satisfies the conditions of Assumption 1: it is compact, convex, $0 \in K_i$ and $K_i \subseteq \mathbb{R}_+^m$. We first consider the loss

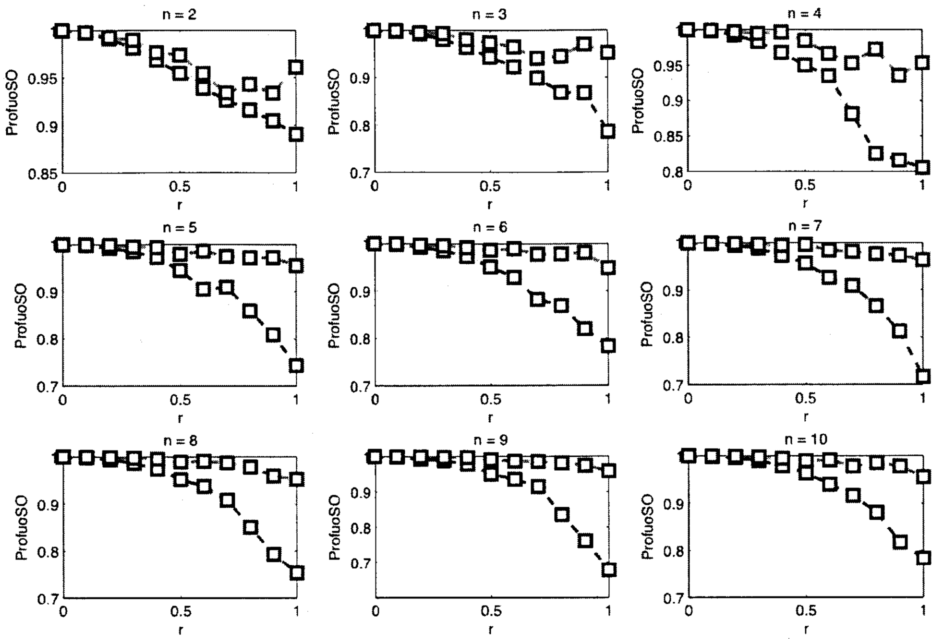


Figure 2-2: Numerical simulation of $\Pi(OP)/\Pi(MP)$ for general constrained (in purple) and unconstrained (in blue) markets.

of total surplus, then at the loss of profit. Our bounds are looser than in the previous section and we do not show tightness since they apply for very general constraints. Still, they provide us with significant guarantees for the loss of surplus and profit. Moreover, as functions of $n$ and $r$, these bounds follow similar trends than the bounds in the previous section.

### 2.4.1 Loss of total surplus

**Theorem 2.5.** *Under Assumptions 1 - 15 for gross substitute products we have:*

$$TS(OP) \geq \frac{5}{6} TS(MP) \tag{2.6}$$

$$TS(MP) \geq \frac{2}{3} TS(SP) \tag{2.7}$$

$$TS(OP) \geq \frac{5}{9} TS(SP) \tag{2.8}$$

As in the case of Section 2.3, where constraints were production capacities, also in the presence of general constraints total surplus bounds are independent of the number of firms $n$ or the market power of the firms $r$. Even in a market when firms are competing and are faced with various convex constraints, the loss of surplus for society is never more than 44% $(= 1 - \frac{5}{9})$. For all values of $n$ and $r$, tightness in Theorem 2.2 and 2.3 shows that a loss of surplus of 25% can be reached (independent firms, unconstrained market) and Theorem 2.5 guarantees that this loss of surplus can never exceed 44%.

Moreover, the oligopoly versus monopoly bound above suggests that under some constraints on the production quantities, allowing firms to compete decreases total surplus. The bound shows that even if competition can in some case hurt the surplus of society, it will at most reduce it by 1/6. We provide below an example where $TS(OP) < TS(MP)$. Our example is the result of random simulations, it is not a parametric example. Also, we relaxed the assumption $\bar{\mathbf{d}} > 0$ to $\bar{\mathbf{p}} > 0$ since the proof below doesn't use this assumption. The example we provide still makes sense as $\mathbf{d}^{OP}$, $\mathbf{d}^{MP}$, $\mathbf{p}^{OP}$ and $\mathbf{p}^{MP}$ are all non-negative. Our example is defined by:

- 2 firms, each selling 2 products ($i = 2$, $m = 2$)
- $r = 0.9$, $\mathbf{M} = \begin{pmatrix} 1.1565 & -0.4947 & -0.1672 & -0.3790 \\ -0.4947 & 2.2010 & -0.4954 & -0.7078 \\ -0.1672 & -0.4954 & 3.0273 & -0.4844 \\ -0.3790 & -0.7078 & -0.4844 & 4.6615 \end{pmatrix}$
- $\bar{\mathbf{p}} = (0.2187, 0.1831, 0.3280, 0.3366)^T$ and $\bar{\mathbf{d}} = \mathbf{M}\bar{\mathbf{p}}$
- Polyhedral production constraints $\mathbf{Ad} \leq \mathbf{b}$ where $\mathbf{A} = \begin{pmatrix} 0.1956 & -0.4221 & 0 & 0 \\ 0 & 0 & -0.8521 & -0.1000 \end{pmatrix}$ and $\mathbf{b} = (0.7214, 0.9625)^T$
- The solutions $\mathbf{d}^{OP}$ and $\mathbf{d}^{MP}$ are computed using constrained quadratic programming (the monopoly problem is a direct quadratic program, and the oligopoly problem can be reformulated into one)

- In this market, $TS(OP) = 0.2258$ which is less than $TS(MP) = 0.2261$

*Proof.* (Theorem 2.5)

At a Nash equilibrium solution, the optimization problem that a single firm faces is:

$$\max_{\mathbf{d}_i} \ \mathbf{d}_i \cdot \left\{ \bar{\mathbf{p}}_i - \begin{pmatrix} B_{i1} \\ \vdots \\ B_{im} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{d}_i \\ \mathbf{d}_{-i}^{OP} \end{pmatrix} \right\}$$

$$\text{s.t.} \qquad \mathbf{d}_i \in K_i$$

where $(B_{ij})_{j=1,\cdots,m}$ denotes the rows of matrix $\mathbf{B}$ corresponding to the different products sold by firm $i$.

Let's denote by $\mathbf{B^{Bdiag}}$ the block diagonal matrix:

$$\mathbf{B^{Bdiag}} = \begin{pmatrix} B_1^1 & 0 & \dots & 0 \\ 0 & B_2^2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & B_n^n \end{pmatrix}$$

where $B_i^i$ denotes the $m * m$ square-matrix for all products of firm $i$.

The variational inequality[2] satisfied at the oligopoly solution is:

$$\left\{ -\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \mathbf{B^{Bdiag}} \cdot \mathbf{d}^{OP} \right\}^T (\mathbf{d} - \mathbf{d}^{OP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Since by definition the monopoly solution must be feasible as well (i.e. $\mathbf{d}^{MP} \in K$), we have:

$$\left\{ -\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \mathbf{B^{Bdiag}} \cdot \mathbf{d}^{OP} \right\}^T (\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0$$

Developing all the terms and extracting $TS(OP) = (\mathbf{B}\bar{\mathbf{d}})^T \, \mathbf{d}^{OP} - 1/2 \, (\mathbf{d}^{OP})^T \mathbf{B} \, \mathbf{d}^{OP}$

[2] A variational inequality VI(F,K) is the problem of finding: $\{x^* \in K, \ F(x^*)^T (x - x^*) \geq 0, \ \forall x \in K\}$. We explain the derivation of this variational inequality in Appendix A.5.

and $TS(MP) = (\mathbf{B}\bar{\mathbf{d}})^T\ \mathbf{d}^{MP} - 1/2\ (\mathbf{d}^{MP})^T\mathbf{B}\ \mathbf{d}^{MP}$, we obtain:

$$\begin{aligned} &TS(OP) - TS(MP) - 1/2\ (\mathbf{d}^{OP})^T\mathbf{B}\mathbf{d}^{OP} - (\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{OP} \\ &\quad -1/2\ (\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP} + (\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{OP} + (\mathbf{d}^{MP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{OP} \geq 0 \end{aligned} \tag{2.9}$$

Using positive definiteness of matrix $\mathbf{B}$, we have:

- $(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{OP} \leq 1/2(\mathbf{d}^{OP})^T\mathbf{B}\mathbf{d}^{OP} + 1/2(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP}$
  $(1/2\ (\mathbf{d}^{OP} - \mathbf{d}^{MP})^T\ \mathbf{B}(\mathbf{d}^{OP} - \mathbf{d}^{MP}) \geq 0)$
- $(\mathbf{d}^{MP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{OP} \leq (\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{OP} + 1/4(\mathbf{d}^{MP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP}$
  $((\mathbf{d}^{OP} - \frac{1}{2}\mathbf{d}^{MP})^T\ \mathbf{B}^{\mathbf{Bdiag}}(\mathbf{d}^{OP} - \frac{1}{2}\mathbf{d}^{MP}) \geq 0)$

Hence, using these two inequalities in (B.16), we get:

$$TS(OP) - TS(MP) + 1/4(\mathbf{d}^{MP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP} \geq 0 \tag{2.10}$$

On the other hand, the variational inequality for the monopoly is:

$$\{-\bar{\mathbf{p}} + \mathbf{B}\cdot\mathbf{d}^{MP} + \mathbf{B}\cdot\mathbf{d}^{MP}\}^T\,(\mathbf{d} - \mathbf{d}^{MP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Evaluating this variational inequality at the feasible vector $0 \in K$ , we have:

$$\bar{\mathbf{p}}\cdot\mathbf{d}^{MP} - 1/2\,(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP} - 3/2\,(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP} \geq 0$$

This leads to:

$$TS(MP) \geq 3/2\,(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP} \Leftrightarrow 1/6\ TS(MP) \geq 1/4\,(\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP}$$

For gross substitute products, matrix $\mathbf{B}$ is an inverse M-matrix with all coefficients non negative so that: $(\mathbf{d}^{MP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP} \leq (\mathbf{d}^{MP})^T\mathbf{B}\mathbf{d}^{MP}$. Using (B.18) and this inequality, (B.17) becomes:

$$\begin{aligned} & TS(OP) - TS(MP) + 1/6\, TS(MP) \geq 0 \\ \Rightarrow\ & TS(OP) \geq 5/6\ TS(MP) \end{aligned}$$

This establishes the desired relation between TS(OP) and TS(MP). Let us now turn to the study of $TS(SP)$. The optimization of the total surplus corresponds to the problem:

$$\max_{\mathbf{d}}\ \bar{\mathbf{p}} \cdot \mathbf{d} - 1/2\ \mathbf{d}^T \mathbf{B}\ \mathbf{d}$$

The optimal solution of this problem satisfies the variational inequality:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{SP}\}^T (\mathbf{d} - \mathbf{d}^{SP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Evaluating this variational inequality at the feasible point 0, we have:

$$TS(SP) \geq 1/2\, (\mathbf{d}^{SP})^T \mathbf{B} \mathbf{d}^{SP} \tag{2.11}$$

On the other hand, the monopoly variational inequality evaluated at the feasible solution $\mathbf{d}^{SP}$ gives:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{MP} + \mathbf{B} \cdot \mathbf{d}^{MP}\}^T (\mathbf{d}^{SP} - \mathbf{d}^{MP}) \geq 0$$

This leads to:

$$TS(MP) - 3/2\, (\mathbf{d}^{MP})^T \mathbf{B} \mathbf{d}^{MP} - TS(SP) - 1/2\, (\mathbf{d}^{SP})^T \mathbf{B} \mathbf{d}^{SP} + 2\, (\mathbf{d}^{SP})^T \mathbf{B} \mathbf{d}^{MP} \geq 0 \tag{2.12}$$

By positive definiteness of matrix **B**, we have:

$$2\, (\mathbf{d}^{SP})^T \mathbf{B} \mathbf{d}^{MP} \leq 3/2\, (\mathbf{d}^{MP})^T \mathbf{B} \mathbf{d}^{MP} + 2/3\, (\mathbf{d}^{SP})^T \mathbf{B} \mathbf{d}^{SP}$$

Making use of this property in (B.19), we finally obtain:

$$
\begin{aligned}
& TS(MP) - TS(SP) + 1/6\,(\mathbf{d}^{SP})^T\mathbf{B}\mathbf{d}^{SP} \geq 0 \\
\Rightarrow\ & TS(MP) - TS(SP) + 1/3\,TS(SP) \geq 0 \quad \text{(by (B.18))} \\
\Rightarrow\ & TS(MP) \geq 2/3\,TS(SP)
\end{aligned}
$$

Finally, combining $TS(OP) \geq 5/6\,TS(MP)$ and $TS(MP) \geq 2/3\,TS(SP)$, we obtain the desired property:

$$TS(OP) \geq 5/9\,TS(SP)$$

$\square$

Similar results hold true for non-linear demand case.

**Theorem 2.6.** *Under Assumptions 1, 14, 6 - 9 for gross substitute products and with similarity coefficient A we have:*

$$TS(OP) \geq \frac{1/A + 1 - 3A/4}{1/2 + A}\, TS(MP) \tag{2.13}$$

$$TS(MP) \geq \left(\frac{5}{3} - A^2\right)\, TS(SP) \tag{2.14}$$

$$TS(OP) \geq \frac{1/A + 1 - 3A/4}{1/2 + A}\left(\frac{5}{3} - A^2\right)\, TS(SP) \tag{2.15}$$

Note in particular that when $A = 1$ as in the linear case, the bounds of Theorem 2.6 reduce to those of Theorem 2.5.

*Proof.* The proof method is similar to that of the linear case. We only sketch the proof here, highlighting the main differences with the previous proof.

At a Nash equilibrium solution, the optimization problem that a single firm faces is now:

$$
\begin{aligned}
\max_{\mathbf{d}_i} \quad & \mathbf{d}_i \cdot P_i(\mathbf{d}) - \mathbf{c}_i\mathbf{d}_i \\
\text{s.t.} \quad & \mathbf{d}_i \in K_i
\end{aligned}
$$

The variational inequality satisfied at the oligopoly solution is:

$$\left\{-P(\mathbf{d}^{OP}) - \mathbf{J}^{\mathbf{diag}}P(\mathbf{d}^{OP})\ \mathbf{d}^{OP} + c\right\}^T (\mathbf{d} - \mathbf{d}^{OP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Similarly, the monopoly problem is:

$$\max_{\mathbf{d}_i} \quad \mathbf{d} \cdot P(\mathbf{d}) - \mathbf{c}\ \mathbf{d}$$
$$\text{s.t.} \qquad \mathbf{d} \in K$$

The corresponding variational inequality is:

$$\{-P(\mathbf{d}^{MP}) - \mathbf{J}P(\mathbf{d}^{MP})\ \mathbf{d}^{MP} + c\}^T (\mathbf{d} - \mathbf{d}^{MP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Finally, optimizing total surplus corresponds to:

$$\max_{\mathbf{d}_i} \quad U(\mathbf{d}) - \mathbf{c}\ \mathbf{d}$$
$$\text{s.t.} \qquad \mathbf{d} \in K$$

The corresponding variational inequality is:

$$\{-P(\mathbf{d}^{SP}) + c\}^T (\mathbf{d} - \mathbf{d}^{MP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

The key idea to connect the variational inequality to $TS(\mathbf{d})$ is to realize that: $P(\mathbf{d}) - \mathbf{c} = \nabla TS(\mathbf{d})$. Using integration by parts, this leads to:

$$TS(\mathbf{d}) = \int_0^{\mathbf{d}} (P(\mathbf{q}) - c)^T \cdot d\mathbf{q} = (P(\mathbf{d}) - c)^T \cdot \mathbf{d} - \int_0^{\mathbf{d}} \mathbf{q} \cdot \mathbf{J}P(\mathbf{q}) \cdot d\mathbf{q}$$

Since $P(\mathbf{d})$ is a gradient field (the gradient of $U(\mathbf{d})$), the integrals above are independent of the path chosen between 0 and $\mathbf{d}$. Using the parametrization $\mathbf{q} = \frac{\mathbf{d}}{T}\ t$

with $t \in [0, T]$, we can rewrite the total surplus as:

$$TS(\mathbf{d}) = \Pi(\mathbf{d}) - \int_0^{\mathbf{d}} \frac{\mathbf{d}}{T} \cdot \mathbf{J}P(\frac{\mathbf{d}}{T}t) \cdot \frac{\mathbf{d}}{T}\; t\; dt$$

Making use of the matrix similarity property:

$$\frac{1}{A}\frac{\mathbf{d}}{T} \cdot (-\mathbf{J}P(\mathbf{d})) \cdot \frac{\mathbf{d}}{T} \leq \frac{\mathbf{d}}{T} \cdot (-\mathbf{J}P(\frac{\mathbf{d}}{T}t)) \cdot \frac{\mathbf{d}}{T} \leq A\,\frac{\mathbf{d}}{T} \cdot (-\mathbf{J}P(\mathbf{d})) \cdot \frac{\mathbf{d}}{T}$$

we can bound $TS(\mathbf{d})$ with the profit $P(\mathbf{d})$

$$\Pi(\mathbf{d}) - \frac{1}{2A}\mathbf{d} \cdot \mathbf{J}P(\mathbf{d}) \cdot \mathbf{d} \leq TS(\mathbf{d}) \leq \Pi(\mathbf{d}) - \frac{A}{2}\mathbf{d} \cdot \mathbf{J}P(\mathbf{d}) \cdot \mathbf{d}$$

Thanks to the concavity of the functions $p_i$ (Assumption 9), we can also compare the price vectors at different demands:

$$P(\widetilde{\mathbf{d}}) \leq P(\mathbf{d}) + \mathbf{J}P(\mathbf{d}) \cdot (\widetilde{\mathbf{d}} - \mathbf{d})$$

Combining these results with the positive definiteness of $-\mathbf{J}P(\mathbf{d})$ (Assumption 8) and carrying out the same steps as in the previous proof, leads to the desired bounds. □

### 2.4.2 Loss of firms profit

**Theorem 2.7.** *For gross substitute products, under Assumptions 1, 14 - 15:*

$$\Pi(OP)/\Pi(MP) \geq \max\left\{\frac{2}{2 + r \cdot (nm-1)}; \frac{3}{4 + r \cdot (nm-1)}\right\}$$

*where $r$ is the market power, $n$ is the number of firms, and $m$ is the number of products.*

The bound is composed of two parts (see Figure 2-3). The first part dominates when $r \cdot (nm - 1) \leq 2$, the second part dominates otherwise.

This bound as well as the tight profit bound of Section 2.3 are decreasing with $n$, $m$ and $r$. Therefore, when more firms or more products participate in the market, the outcome of selfish behavior becomes less efficient in terms of total profit. With $r = 1$, the profit ratio converges to 0 as $nm$ goes to $\infty$, meaning the loss of profit due to competition can become arbitrarily bad when the number of firms or the number of products is large. On the contrary, in the monopoly setting of a single firm selling a single product $nm = 1$, the profit ratio is one because there is no loss of profit. Similarly when $r$ equals 0, every firm is operating as a monopolist because the price set by one firm does not affect the demand of its competitors. Hence there is no inefficiency in terms of profits and the profit ratio is one. As competition intensifies (r gets closer to 1), the oligopoly solution becomes more "inefficient".

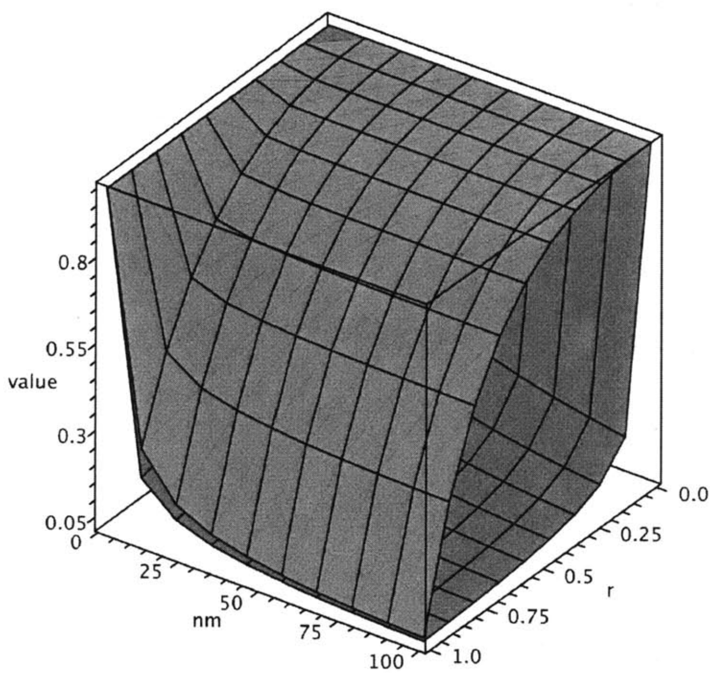


Figure 2-3: Bounds on the profit ratio $\Pi(OP)/\Pi(MP)$ for markets with simple capacity constraints (green) and general convex constraints (red).

In order to prove the previous theorem, we will need to use an intermediate inequality. To shorten notations, we denote by $\|\mathbf{d}\|_{\mathbf{B}}^2 = \mathbf{d}^T\ \mathbf{B}\mathbf{d}$, $\|\mathbf{d}\|_{\mathbf{B}^{\mathbf{Bdiag}}}^2 = \mathbf{d}^T\ \mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}$, $\|\mathbf{d}\|_{\mathbf{\Gamma}}^2 = \mathbf{d}^T\ \mathbf{\Gamma}\mathbf{d}$.

**Lemma 2.2.** *For a symmetric inverse M-matrix* $\mathbf{B}$ *and a vector* $\mathbf{d}$ *with all component positive, the following inequality holds:*

$$\|\mathbf{d}\|_{\mathbf{B}}^2 \leq (1 + r \cdot (nm - 1))\ \|\mathbf{d}\|_{\mathbf{B}^{\mathbf{Bdiag}}}^2$$

*where* $r$ *is the market power.*

The proof of this lemma can be found in appendix A.4. We are now ready to prove the main theorem:

*Proof.* The variational inequality satisfied at the oligopoly solution is:

$$\left\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \mathbf{B}^{\mathbf{Bdiag}} \cdot \mathbf{d}^{OP}\right\}^T (\mathbf{d} - \mathbf{d}^{OP}) \geq 0 \qquad \forall \mathbf{d} \in K \tag{2.16}$$

Since by definition the monopoly solution must be feasible as well (i.e. $\mathbf{d}^{MP} \in K$), the first order optimality conditions for the monopoly, become equivalent to:

$$\left\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \mathbf{B}^{\mathbf{Bdiag}} \cdot \mathbf{d}^{OP}\right\}^T (\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0$$

Denoting by $\Pi = \mathbf{d} \cdot \{\bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}\}$ the profit, we get:

$$\Pi(OP) - \|\mathbf{d}^{OP}\|_{\mathbf{B}^{\mathbf{Bdiag}}}^2 - \bar{\mathbf{p}}^T\mathbf{d}^{MP} + (\mathbf{d}^{OP})^T\mathbf{B}\mathbf{d}^{MP} + (\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP} \geq 0$$

By adding and substracting $(\mathbf{d}^{MP})^T B\mathbf{d}^{MP}$, we obtain:

$$\Pi(OP) - \Pi(MP) + (\mathbf{d}^{OP})^T\mathbf{B}\mathbf{d}^{MP} - \|\mathbf{d}^{MP}\|_{\mathbf{B}}^2 - \|\mathbf{d}^{OP}\|_{\mathbf{B}^{\mathbf{Bdiag}}}^2 + (\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP} \geq 0 \tag{2.17}$$

<u>Part 1:</u> Let's first prove the second part of the bound, namely $\Pi(OP)/\Pi(MP) \geq \frac{3}{4+r\cdot(nm-1)}$.

Since both $\mathbf{B}$ and $\mathbf{B}^{\mathbf{Bdiag}}$ are positive definite matrices we can use the following bounds:

- $(\mathbf{d}^{OP})^T\mathbf{B}\mathbf{d}^{MP} \leq 1/3\|\mathbf{d}^{OP}\|_{\mathbf{B}}^2 + 3/4\|\mathbf{d}^{MP}\|_{\mathbf{B}}^2$
- $(\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}\mathbf{d}^{MP} \leq \|\mathbf{d}^{OP}\|_{\mathbf{B^{Bdiag}}}^2 + 1/4\|\mathbf{d}^{MP}\|_{\mathbf{B^{Bdiag}}}^2$

Introducing these bounds into the variational inequality:

$$\Pi(OP) - \Pi(MP) + 1/3\|\mathbf{d}^{OP}\|_{\mathbf{B}}^2 \underbrace{-1/4\|\mathbf{d}^{MP}\|_{\mathbf{B}}^2 + 1/4\|\mathbf{d}^{MP}\|_{\mathbf{B^{Bdiag}}}^2}_{=-1/4\|\mathbf{d}^{MP}\|_{\mathbf{B^{BOffdiag}}}^2 \leq 0} \geq 0$$

Since we assumed $0 \in K$, we can plug the feasible point 0 into the variational inequality (B.1) to get:

$$\Pi(OP) \geq \|\mathbf{d}^{OP}\|_{\mathbf{B^{Bdiag}}}^2 \tag{2.18}$$

Using Lemma B.1 to upper bound $\|\mathbf{d}^{OP}\|_{\mathbf{B}}^2$, we finally get:

$$\begin{aligned}\Pi(OP) - \Pi(MP) + 1/3(1 + r \cdot (nm-1))\ \|\mathbf{d}^{OP}\|_{\mathbf{B^{Bdiag}}}^2 &\geq 0\\ \Pi(OP) - \Pi(MP) + 1/3(1 + r \cdot (nm-1))\ \Pi(OP) &\geq 0\end{aligned}$$

This inequality is equivalent to:

$$\frac{\Pi(OP)}{\Pi(MP)} \geq \frac{3}{4 + r \cdot (nm-1)}$$

<u>Part 2:</u> Let's now prove the first part of the bound $\Pi(OP)/\Pi(MP) \geq \frac{2}{2+r\cdot(nm-1)}$.

Using symmetry of the matrix $\mathbf{B}$, we can rewrite equation (B.2) as follow:

$$\Pi(OP) - \Pi(MP) + (\mathbf{d}^{OP})^T\mathbf{B}^{\mathbf{Bdiag}}(\mathbf{d}^{MP} - \mathbf{d}^{OP}) - (\mathbf{d}^{MP})^T\mathbf{B}(\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0$$

We can decompose this expression in two different ways:

- $$\Pi(OP) - \Pi(MP) + \underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{MP})^T \mathbf{B}^{\mathbf{Bdiag}} (\mathbf{d}^{MP} - \mathbf{d}^{OP})}_{\leq 0} - (\mathbf{d}^{MP})^T \mathbf{B}^{\mathbf{BOffdiag}} (\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0 \tag{2.19}$$

- $$\Pi(OP) - \Pi(MP) + \underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{MP})^T \mathbf{B} (\mathbf{d}^{MP} - \mathbf{d}^{OP})}_{\leq 0} - (\mathbf{d}^{OP})^T \mathbf{B}^{\mathbf{BOffdiag}} (\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0 \tag{2.20}$$

Combining these two expressions (1/2 Inequality(B.4) + 1/2 Inequality(B.5)) and leaving out the non-positive terms, we get:

$$\Pi(OP) - \Pi(MP) - 1/2(\mathbf{d}^{MP} + \mathbf{d}^{OP})^T \mathbf{B}^{\mathbf{BOffdiag}} (\mathbf{d}^{MP} - \mathbf{d}^{OP}) \geq 0$$
$$\Pi(OP) - \Pi(MP) \underbrace{-1/2\ (\mathbf{d}^{MP})^T \mathbf{B}^{\mathbf{BOffdiag}} \mathbf{d}^{MP}}_{\leq 0} + 1/2\ (\mathbf{d}^{OP})^T \mathbf{B}^{\mathbf{BOffdiag}} \mathbf{d}^{OP} \geq 0$$

Using Lemma B.1, we can write:

$$\|\mathbf{d}^{OP}\|^2_{\mathbf{B}^{\mathbf{BOffdiag}}} = \|\mathbf{d}^{OP}\|^2_{\mathbf{B}} - \|\mathbf{d}^{OP}\|^2_{\mathbf{B}^{\mathbf{Bdiag}}} \leq r \cdot (nm-1) \|\mathbf{d}^{OP}\|^2_{\mathbf{B}^{\mathbf{Bdiag}}}$$

With inequality (B.3), we finally obtain:

$$\Pi(OP) - \Pi(MP) + 1/2\ r \cdot (nm-1)\Pi(OP) \geq 0$$

This leads to:

$$\frac{\Pi(OP)}{\Pi(MP)} \geq \frac{2}{2 + r \cdot (nm-1)}$$

□

Here as well the result extends to the non-linear case. The bound is actually the same for linear and non-linear demand. Since the profit bound does not entail calculations of total surplus, we do not even need the similarity property and as a consequence $A$ does not appear in the bound.

**Theorem 2.8.** *For gross substitute products, under Assumptions 1, 14, 6 - 9:*

$$\Pi(OP)/\Pi(MP) \geq \max\left\{\frac{2}{2+r\cdot(nm-1)}; \frac{3}{4+r\cdot(nm-1)}\right\}$$

*where $r$ is the market power, $n$ is the number of firms, and $m$ is the number of products.*

*Proof.* Again, the proof method is similar to that of the linear case. We do not need the similarity property.

We make use of the concavity assumption of $p_i$ (Assumption 9) to write:

$$P(\mathbf{d}^{MP}) \leq P(\mathbf{d}^{OP}) + \mathbf{J}P(\mathbf{d}^{OP}) \cdot (\mathbf{d}^{MP} - \mathbf{d}^{OP})$$

We also take advantage of the positive definiteness of $-\mathbf{J}P(\mathbf{d})$ (Assumption 8). Carrying out similar steps as in the linear case, we obtain the desired result. □

## 2.5 Conclusion

We studied the effects of competition in a market where firms compete by deciding production quantities. We looked at the loss of surplus for society and the loss of firms' profit. For the case of single-product firms facing simple capacity constraints, we derived tight bounds for the worst-case losses as a function of the number of firms $n$ and the intensity of competition as measured by $r$. As it turns out, the maximum loss of surplus is 25% independently of $n$ or $r$. The maximum loss of profit on the other hand, is a decreasing function of $n$ and $r$. Even when competition is fierce ($r = 1$), if the number of competing firms is small, our bound guarantees a small loss of profit (e.g. for 2 firms, the loss of profit never exceeds 13%). Moreover, even for an arbitrarily large number of firms, if $r < 1$, our bound still provides a significant guarantee (see Figure 2-3). We then extended our model to multi-product firms facing general convex production constraints. For this general case, we were able to produce total surplus and profit bounds with a flavor similar to the previous ones.

We guarantee a loss of surplus smaller than 44% and we exhibit a bound for the loss of profit as a function of $n$, the number of product $m$, and $r$. Although tightness is not shown, together, the simple capacity bounds and the general constraint bounds provide a range where the maximum loss of surplus and profit lie.

# Chapter 3

# Enforcing a joint energy consumption target on companies with multiple subsidiaries: a decentralized incentive mechanism and its social impact

## 3.1 Introduction

### 3.1.1 Motivation

In response to increasing social pressure, more and more companies, especially large corporations, are committed to energy consumption reduction goals as part of their Corporate Social Responsibility (CSR) initiative. PepsiCo, for instance, has pledged to cut water consumption by 20 percent, electricity by 20 percent and fuel by 25 percent by 2015. P&G also announced plans to cut energy and water use by 20 percent by 2012 (see [108]).

In the long run, these consumption reduction targets encourage technology innovation, and ultimately increase energy efficiency. In the short run, however, since output level is often directly linked to energy consumption, the limitations in energy consumption effectively place a cap on the output level. This is often then translated into a series of production quotas for subsidiaries or divisions within the parent company. This gives rise to the typical 'principal-agent' problem where divisional managers deliberately overlook their assigned quotas in order to achieve a higher divisional profit, especially when the burden of over-consumption is not borne by each individual division or subsidiary.

Since company-wide coordination is often operationally impossible or too costly, it is necessary for the parent company to design alternative mechanisms to induce the 'optimal' behavior of its divisional managers. That is, a total consumption level under the promised energy reduction goal, and a distribution among divisions that maximizes total profit for the company. One common approach to align the different incentives is to use financial compensation. For example, since 2008, Intel has established an employee engagement mechanism where a portion of each employee's variable compensation is dependent upon the company achieving its environmental sustainability goals [70]. Similarly, Hilton Hotels successfully met its goal of 5 percent reduction in energy consumption by tying the hotel general managers' annual bonuses to the energy performance at each property [59].

This chapter models a company attempting to reduce its global energy consumption across multiple subsidiaries. The subsidiaries behave like Cournot players, who independently decide their production quantity to optimize individual profit. Cournot competition (see [33]) is often a good model of competition in industries where output level cannot be easily adjusted in the short-run (the manufacturing industry for example). In a Cournot market, decisions on output quantities are made first and prices are determined subsequently to clear the market. In this competitive setting, the demand of a particular division is adversely affected by the output of other divisions; this could happen when divisions sell the same or slightly differentiated products in the same market, (i.e. different brands of shampoo from P&G). In the absence of any coordination mechanism, the subsidiaries compete as Cournot players subject to the joint energy target constraint. Each subsidiary is conscious of the combined energy constraint the parent company must satisfy: it chooses its output to maximize profit subject to the satisfaction of the energy constraint while taking the production quantities of its competitors as given. We call this situation the **decentralized scenario**; equilibrium is reached when no subsidiary can increase its profit by unilaterally changing its output quantity. Compared to problems with disjoint constraints, where the feasible set of each subsidiary is independent of the decisions of its competitors, the equilibrium behavior is more complicated under joint constraints since subsidiaries influence each other through both their objective function and their feasible strategy space; in particular, such a game possesses multiple equilibria. Another option for the company is to coordinate the efforts of its subsidiaries by subsidizing reduced energy consumption. Under such a **subsidy system**, each subsidiary is free to decide its output quantity without worrying about the energy constraint. It is up to the parent company to choose an appropriate level of subsidies so that it is optimal for the subsidiaries to consume less than the energy target. Finally, the company also has the option to **centrally** determine the optimal production and energy consumption levels of its subsidiaries and to then impose these levels on them. This chapter proposes an **implementation of a reward mechanism** (a subsidy system) for the parent company to coordinate its subsidiaries. It then **compares the subsidy system** with the fully decentralized Cournot scenario and the centrally-controlled scenario in terms of company profit and social welfare. The chapter shows that **there always exist an optimal subsidy system** that is able to fully coordinate the subsidiaries and achieve maximum company profit. It also describes an **operationally simpler subsidy system** that achieves close-to-optimal company profit in a number of situations. The chapter finally studies the impact of the joint energy target. This constraint makes the analysis much harder because the Cournot Nash equilibria are now the solutions of QVI's (quasi variational inequalities, see [11]) instead of variational inequalities. The joint energy constraint radically affects the Cournot model. Without the joint constraint, the Cournot model admits a unique equilibrium solution and the loss of profit and social welfare resulting from free competition are bounded. In the presence of a joint constraint, there exists a set of Nash equilibria (no more uniqueness) and the loss of profit and social welfare resulting from Cournot competition can, in general, be arbitrarily bad. The chapter analyzes how the joint energy constraint affects these losses.

### 3.1.2 Literature review

As President Nixon put it in his 1970 State of the Union Address: "We can no longer afford to consider air and water common property, free to be abused by anyone without regard to the consequences. Instead, we should begin now to treat them as scarce resources which we are no more free to contaminate than we are free to throw garbage into our neighbors yard." Over the last thirty years, governments and international organizations have been drafting regulations to limit the amount of pollution generated by industrial activities. As a consequence, lawmakers and environmental economists alike have been studying a variety of regulation tools to efficiently limit pollution without impacting the economy more than necessary. Four main types of policy instruments have emerged in the literature: i) quotas, ii) Pigouvian taxes, iii) ambient taxes and iv) pollution permits. The most direct way to limit pollution is the use of **quotas** which determine the maximum amount of pollutant a given industry, region or factory is allowed to emit. Another way to reduce pollution is to tax the polluting activity. In an ideal world with perfect competition, perfect information of the regulator and no pre-existing taxes, quotas and taxes are substitutable tools. In practice however, taxes are usually preferable to quotas because they allow the regulator to achieve a socially optimal configuration even under asymmetric information between the polluters and the regulator (Weitzman [137], Hoel & Karp [65]) or with preexisting distortionary taxes (Parry [100]). There are in fact two kinds of taxes used as regulatory instruments. The well known **Pigouvian tax** considers the damage generated by pollution as an externality, not accounted for by the polluting companies, and recommends taxing it at a rate equal to the marginal social damage caused by pollution. This procedure forces firms to internalize the cost of their polluting activity and is hence able to achieve Pareto efficiency (Wellisz [138], Wright [141]). This tax assumes that the regulator is able to monitor and charge each firm on the basis of its individual emission. This kind of tax is called a point-tax, as opposed to an ambient tax which taxes firms in an industry on the basis of their global emission (the sum of their emissions). Under an **ambient tax**, each firm pays the same amount based on the total emission of the industry. A major drawback of such a tax is its potential to cause large transfers of wealth between firms. Yet, in many cases where the pollution of each firm cannot be measured individually, the ambient tax provides an implementable policy. Moreover, Karp [73] shows that in an oligopoly industry, each firm might end up paying less taxes under an ambient tax than under a point tax: this is because, under an ambient tax, each firm realizes its impact on the global pollution level. The last tool available to the regulator is to issue a limited amount of tradable **pollution permits** to the firms. A firm can only pollute up to the amount of pollution licenses it owns; this limits the global pollution level to the amount of permits issued by the regulator. The regulator can either allocate them for free (grandfather them) at the beginning of the period or auction them. Montgomery [87] demonstrates that markets in licenses achieve maximum profit for the firms and Nagurney & Dhanda [91] show that markets in licenses have a well-defined, unique and easily computable equilibrium. Papers on markets in pollution licenses (including the two papers above) all share a common limitation though. To reflect reality, they model industry as an oligopoly of firms, yet they assume the market in pollution licenses to be perfectly competitive. The underlying reason for this limitation is that when firms are price takers for the pollution permits, their joint pollution constraint decouples allowing the market to be modeled through a variational inequality instead of the much harder quasi variational inequality framework. There is no fundamental justification though why firms in an oligopoly market would be price takers of the pollution rights.

In this context, we propose a new regulation instrument that avoids the drawbacks of existing policies. Since current models of tax systems outperform both quotas and permits by providing a simpler, more practical and more flexible tool, our instrument is based on taxes. The main limitation of the Pigouvian literature is the assumption that the social damage of pollution is easily evaluated in financial terms. What is the exact monetary cost of green house gas emissions or water pollution? Instead our model sets a global energy consumption (or pollution) standard, a limit based on environmental considerations, and then puts in place a system of taxes to reach the standard while minimizing the loss of social welfare for firms and consumers. In fact, the idea of such a system has been toyed with before in the environmental literature. Baumol & Oates [10] point out several limitations of Pigouvian taxes and suggests the use of what they call "standards and prices". However, they do not develop a mathematical model of the tool or analyze its financial implications. This is probably because of the complexity of the quasi-variational inequality framework. The goal of this chapter is precisely to develop the mathematical analysis of the standards and price model, to compute optimal taxes and to bound the loss of efficiency resulting from alternative simpler tax rules.

In a broader sense, our model is an attempt to coordinate a horizontal supply chain under a joint capacity constraint through the use of taxes set by a regulator. As such, this chapter is part of the supply chain coordination literature. Excellent reviews on the topic can be found in, for example, Lariviere [81], Tsay et al. [129] and Cachon [24]. A large part of this literature is devoted to coordinating a vertical supply chain composed of one supplier and one retailer under wholesale-price, buy-back, revenue-sharing and other types of contracts (Pasternack [101], Bernstein & Federgruen [14], Lariviere & Porteus [82], Cachon & Lariviere [25]). Closer to our model, contracts to coordinate a supply chain with one supplier and multiple competing retailers have been studied in a number of papers (Padmanabhan & Png [97], Deneckere et al. [109], Cachon & Lariviere [25], Bernstein & Federgruen [15]). However, the supply chain they analyze differs from ours in several aspects. First, our model does not focus on supply chain coordination between a supplier and its retailers. Moreover, our measure of efficiency accounts for consumer surplus. We are not only interested in the profit of the supply chain; we are looking for a regulation that is socially efficient. There are some papers studying the loss of social welfare or "price of anarchy" (a term first used in Koutsoupias & Papadimitriou [77]) under horizontal competition without a supplier (see also Bernstein & Federgruen [12], Guo & Yang [60], Farahat & Perakis [46], Kluberg & Perakis [76]). The crucial difference between those papers and this one is the presence of a joint energy consumption target decided by the regulator (see previous paragraph for its justification). This energy target places a hard constraint on the joint productions of the subsidiaries. While competition in the unconstrained (or uncoupled constrained) models leads to a unique Nash equilibrium (where no firm can improve its profit by unilaterally changing its strategy), competition under 'joint' or 'coupled' constraints belongs to the class of generalized Nash equilibrium problems (GNEPs). Our improved regulation instrument comes at the price of a more complex analysis.

The concept of generalized Nash equilibrium (GNE) (often referred to as pseudo-game or abstract economy) was first formally introduced in Debreu [39]. It is a generalization of the Nash equilibrium concept where the choice of an action by one player affects both the pay-off and the domain of action of other players. Unlike the class of Nash equilibrium problems, uniqueness of solution is rare for GNEPs. Rosen [113] introduced the notion of the normalized Nash equilibrium as a special kind of generalized Nash equilibrium and provided conditions on its existence and uniqueness. We will define and analyze the properties of the normalized Nash equilibrium in our environmental game. GNEPs have a wide application in modeling problems where a common resource is shared by players. Pang & Fukushima [98] explore GNEPs in multi-leader follower games; they focus on existence and solution algorithms. Pang et al. [99] study an application to power allocation problems in telecommunications, and Adida & Perakis [1] investigate a model in dynamic pricing and inventory management. We refer the reader to Facchinei & Kanzow [45] for a comprehensive survey on GNEPs.

A few papers apply the framework of generalized Nash equilibria to environmental problems. Haurie & Krawczyk [61] analyze a river pollution problem with a mix of oligopolistic and price taking participants. The chapter exhibits a normalized Nash equilibrium of the game and a system of taxes to enforce it. Contrary to ours, their paper does not worry about the multiplicity of normalized Nash equilibria (and corresponding taxes) or about the performance of the equilibrium in terms of firms profit and social welfare. More recently, Krawczyk [78] presents a model with a global pollution standard and establishes conditions for existence and uniqueness of an equilibrium (normalized Nash equilibrium). The author acknowledges the importance of charging optimal taxes to the producers to achieve social efficiency but does not derive these charges in closed form. Instead, their analysis is limited to computational experiments. The chapter most closely related to ours is Tidball and Zaccour [127]. They study an environmental game under three different scenarios: a Nash equilibrium under separable constraints for each firm, normalized equilibria under joint constraints tying all firms together and a fully cooperative solution. The scenarios are compared in terms of firms profit and total emission level. The chapter focuses mainly on a two-player game. In contrast, this chapter studies an n-player game. It evaluates the scenarios in terms of the welfare of society (including consumers' surplus), provides bounds on the loss of social welfare and firm's profit (rather than simply stating which scenario performs best), and analyzes equilibrium prices and quantities.

### 3.1.3 Main contributions and chapter outline

This chapter brings three main contributions to the environmental regulation literature.

First, it quantifies the losses of profit and social welfare due to the lack of coordination between subsidiaries of a company subject to a joint energy target. It compares the performance of Cournot decentralized competition, the subsidy system and the centrally-controlled scenario. Since equilibrium strategies are not unique under a constrained Cournot game, we focus on two special equilibria: the worst Nash equilibrium and the uniform Nash equilibrium (to be described in the next section). The losses of profit and social welfare are evaluated as functions of various market characteristics (i.e., number of subsidiaries, intensity of competition and asymmetry between subsidiaries). In addition, the chapter examines the equilibrium strategies themselves and explains the underlying rationale behind the discrepancies in price and quantity.

Second, the chapter proposes a decentralized reward-based mechanism for a company to coordinate the energy reduction effort of its subsidiaries. The chapter shows that an appropriate system of subsidies can fully coordinate the subsidiaries and achieve maximum company profit. In the process, we demonstrate that every equilibrium of the constrained Cournot game can be obtained as the Nash equilibrium of an unconstrained game under the right set of subsidies. The chapter also illustrates that an operationally simple (uniform) reward scheme is able to achieve close-to-optimal company profit when the subsidiaries exhibit a certain degree of symmetry.

Third, the chapter analyzes the impact of the joint energy target constraint on Cournot competition. When the energy constraint is not active, there is a unique Nash equilibrium solution, the profit loss due to deregulation is bounded and the loss of social welfare cannot exceed $1/4$. When the joint constraint is tight, on the other hand, there is a continuum of Nash equilibria (instead of a unique one) and the loss of profit and social welfare can become arbitrarily large even when the company only has two subsidiaries.

The chapter is structured as follows. We first introduce the notations and assumptions. We then present the mathematical formulation of a reward-based incentive mechanism and demonstrate its equivalence to the general model of Cournot competition under a joint constraint. In Section 3.3, we show how the presence of a joint constraint magnifies the loss of profit and social welfare. Sections 3.4 and 3.5 present special cases of the problem with various market characteristics. Section 3.4

considers a market consisting of many symmetric subsidiaries: subsidiaries have the same price potential and are uniform in their ability to influence their own selling price as well as the prices of their competitors. We then relax this assumption to consider subsidiaries with only symmetric price potential in Section 3.5. Section 3.6 analyzes the loss of profit and social welfare for both the worst Nash equilibrium and the uniform Nash equilibrium without any restrictions on the market characteristics. We conclude with some insights on the practical implications of these policies.

## 3.2 Model Description

Consider a company with $n$ subsidiaries committed to reducing its global energy consumption to a level $C$.

*Assumption* 10. The subsidiaries sell differentiated substitute products to the market.

Since we are interested in the effects of competition within the company, we will assume the state of the economy outside the company as given. In particular, the quantities and prices set by producers of similar products outside the company are viewed as fixed. Within the company, the subsidiaries compete by deciding their production (and selling) quantities as in the Cournot model. The subsidiary managers choose their output level $d_i$ simultaneously and independently with the objective of maximizing their division's profit. Vector $\mathbf{d} = (d_1, .., d_n)$ denotes the subsidiaries' output levels.

As is traditional in the literature (see [135] Chap. 6), we model customers of this market via a representative consumer. This consumer values the possession of quantities $\mathbf{d}$ of products according to the quadratic utility function:

$$U(\mathbf{d}) = (\widetilde{\mathbf{p}})^T \ \mathbf{d} - 1/2 \ \mathbf{d}^T \mathbf{B} \ \mathbf{d}$$

For now, vector $\widetilde{\mathbf{p}}$ and matrix $\mathbf{B}$ are just parameters of the utility function.

The consumer has to balance the utility he gets from owning these products against the money he spends to buy them. Therefore, for a fixed price vector $\mathbf{p}$, we define the consumer surplus as:

$$CS(\mathbf{d}) = U(\mathbf{d}) - \mathbf{p} \cdot \mathbf{d}$$

To decide the quantities to buy, the consumer maximizes his surplus: $\max_{\mathbf{d}}\ CS(\mathbf{d})$. Hence for a fixed $\mathbf{p}$, this consumer buys quantities $\mathbf{d}$ of products satisfying: $\mathbf{p} = \widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$. This, in turn, gives rise to the affine, invertible demand function:

$$\mathbf{d}(\mathbf{p}) = \widetilde{\mathbf{d}} - \mathbf{M} \cdot \mathbf{p} = \begin{pmatrix} \widetilde{d}_1 \\ \vdots \\ \widetilde{d}_i \\ \vdots \\ \widetilde{d}_n \end{pmatrix} - \begin{pmatrix} M_{11} & -M_{12} & \dots & -M_{1n} \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ -M_{n1} & \dots & -M_{n(n-1)} & M_{nn} \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ \vdots \\ p_i \\ \vdots \\ p_n \end{pmatrix}$$

where matrix $\mathbf{B}$ is the inverse of matrix $\mathbf{M}$.

The subsidiaries aim to sell their entire production output; they set prices to clear the market. The resulting vector of prices $\mathbf{p}_i(d_i, \mathbf{d}_{-i})$ is simply:

$$\mathbf{p}(\mathbf{d}) = \widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$$

$\Gamma$ denotes the diagonal matrix containing the diagonal elements of $\mathbf{B}$. $\widetilde{\mathbf{p}} = \mathbf{B}\ \widetilde{\mathbf{d}}$ is called the price potential since it represents the maximum market prices of the products. The corresponding vector $\widetilde{\mathbf{d}}$ is the demand potential.

We denote by $\mathbf{c}_i(\mathbf{d}_i)$ the vector of unit production costs for subsidiary $i$ and we assume that it only depends on the subsidiaries' own production. The profit of the company can be expressed as:

$$\Pi(\mathbf{d}) = \mathbf{p}(\mathbf{d})^T\mathbf{d} - \mathbf{c}(\mathbf{d})^T\mathbf{d} = \mathbf{d}^T(\bar{\mathbf{p}} - \mathbf{B}\mathbf{d} - \mathbf{c}(\mathbf{d}))$$

Another quantity of interest is total surplus which measures the benefit of a market for society as a whole. Total surplus is defined by aggregating the consumer surplus and the firm's profit:

$$\begin{aligned} TS(\mathbf{d}) &= CS(\mathbf{d}) + \Pi(\mathbf{d}) = U(\mathbf{d}) - \mathbf{p}(\mathbf{d}) \cdot \mathbf{d} + [\mathbf{p}(\mathbf{d}) - \mathbf{c}(\mathbf{d})] \cdot \mathbf{d} \\ &= [\widetilde{\mathbf{p}} - \mathbf{c}(\mathbf{d})]^T \cdot \mathbf{d} - 1/2\ \mathbf{d}^T \mathbf{B}\ \mathbf{d} \end{aligned}$$

*Assumption* 11. We assume that the energy consumption level is linked directly to the output quantities. As a result, the target energy level is modeled as a constraint on the company's output:

$$\mathbf{e}^T \mathbf{d} \leq C.$$

This assumption makes sense as in the short term company cannot drastically improve their pollution efficiency, so that the only way to reduce pollution is to decrease production output.

*Assumption* 12. The demand function is affine: $\mathbf{d}(\mathbf{p}) = \widetilde{\mathbf{d}} - \mathbf{M} \cdot \mathbf{p}$, with $\mathbf{M}$ positive definite, $\text{diag}(\mathbf{M}) > 0$, and $\text{offdiag}(\mathbf{M}) \leq 0$.

As is typical in the literature, function $U(\mathbf{d})$ is strictly concave, which translates into its Hessian matrix $\mathbf{B}$ being positive definite. This leads to $\text{diag}(\mathbf{M}) > 0$. The subsidiaries of the company sell substitute products which means that as more consumers purchase from one subsidiary, less purchase from the others. This implies $\text{offdiag}(\mathbf{M}) \leq 0$

*Assumption* 13. $\mathbf{M}$ is a symmetric, diagonally-dominant matrix.

Symmetry is a consequence of the representative consumer utility assumption. $\mathbf{M}$ is strictly diagonally-dominant means: $M_{ii} > \sum_{j \neq i}^{n} M_{ij}$ for all $i = 1, 2, ...., n$. This assumption is applicable to markets where total demand is decreasing with prices. Under these assumptions, matrix $\mathbf{M}$ belongs to the class of M-matrices. We refer the reader to Horn and Johnson [68] for the definition and properties of M-matrices.

The diversion ratio (see [20] for reference) is defined as:

$$r_i = \frac{\sum_{j \neq i}^{n} \mathbf{M}_{ij}}{\mathbf{M}_{ii}} \in [0, 1), \ \forall i = 1, 2, ..., n, \quad \text{and} \quad r = \max_i r_i \tag{3.1}$$

$r$ measures the intensity of competition since when $r \to 0$, $M$ is a diagonal matrix, and the strategy of a particular subsidiary has little influence on the strategy of other subsidiaries; when $r \to 1$, the total demand in the market remains constant regardless of the market prices, suggesting a highly competitive market where the customers lost by a particular subsidiary following a price increase are captured entirely by its competitors.

*Assumption* 14. We restrict attention to constant per unit production costs $\mathbf{c}$.

Modulo some small technical assumptions, the results of the chapter still hold true for linear per unit costs but it makes notations more complex. We simplify notations by combining unit prices and unit production costs into a unit profit. We denote by $\widetilde{\mathbf{p}}$ the constant term of the price function $\mathbf{p}(\mathbf{d}) = \widetilde{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$ and by $\bar{\mathbf{p}} = \widetilde{\mathbf{p}} - \mathbf{c}$ the constant term of the per-unit profit function $\pi(\mathbf{d}) = \bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}$. The profit of the subsidiaries is thus: $\mathbf{d} \cdot (\bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d})$, their production quantities times their unit profit. Similarly, we introduce $\widetilde{\mathbf{d}} = \mathbf{M} \cdot \widetilde{\mathbf{p}}$ and $\bar{\mathbf{d}} = \mathbf{M} \cdot \bar{\mathbf{p}}$.

*Assumption* 15. We assume $\bar{\mathbf{d}} = \mathbf{d}(\mathbf{c}) > 0$.

This assumption means that every subsidiary can make a non-negative profit by charging just above its production cost.

In the rest of this section, we describe the different approaches available to the company to achieve its reduced energy consumption target. (i) In a completely decentralized approach the company can let its subsidiaries compete freely subject to the satisfaction of the energy target. This leads to Cournot competition subject to a joint production constraint. Due to the presence of the joint constraint, the solutions of this game are generalized Nash equilibria. (ii) The company can also use financial subsidies as an incentive for the subsidiaries to reduce their energy consumption. We show next that every generalized Nash equilibrium of the constrained Cournot game can be achieved in a decentralized manner by the subsidiaries under the right set of subsidies. (iii) We finally define the fully centralized, company optimal and socially optimal, solutions.

### 3.2.1 Cournot competition under a joint constraint

In the absence of a coordination mechanism, each subsidiary $i = 1, \ldots, n$ chooses an output level $d_i^{OP}$ that maximizes its profit subject to the satisfaction of the global energy consumption target assuming the other subsidiaries set their output levels to $\mathbf{d}_{-i}^{OP}$. This target places a constraint on the total output:

$$\begin{aligned} d_i^{OP} = \operatorname{argmax}_{d_i} \Pi(d_i, \mathbf{d}_{-i}^{OP}) \\ \text{s.t. } d_i + \sum_{j \neq i} d_j^{OP} \leq C \\ d_i \geq 0 \end{aligned} \tag{3.2}$$

A simultaneous solution to all the subsidiaries' optimizations is called a generalized Nash equilibrium. At this point, no subsidiary can improve its profit by unilaterally changing its strategy within the feasible strategy space defined by the strategies of the other subsidiaries. For the rest of the chapter, we refer to problem (3.2) as the oligopoly problem (OP), and use $S = \{\mathbf{d}^{OP}\}$ to denote the set of equilibria. Existence of generalized Nash equilibria under very broad assumptions satisfied for the model in this chapter, is established by Debreu [39].

The KKT conditions of the problem are:

$$\begin{cases} \bar{\mathbf{p}} - (\mathbf{B} + \Gamma)\mathbf{d}^{OP} - \boldsymbol{\mu} + \boldsymbol{\lambda} = \mathbf{0} \\ \mathbf{e}^T \cdot \mathbf{d}^{OP} \leq C \\ \mu_i (C - \mathbf{e}^T \cdot \mathbf{d}^{OP}) = 0 \\ \lambda_i \, d_i^{OP} = 0 & \forall i = 1, 2, ..., n \\ \lambda_i \geq 0, \mu_i \geq 0, d_i^{OP} \geq 0 \end{cases} \tag{3.3}$$

The solution to the oligopoly problem is not unique since the $n$ multipliers $\mu_i$ are all associated (through complementarity) to one production constraint. The KKT system is underdetermined: if the production constraint is active, all the $\mu_i$'s are free and there are $3n$ unknowns ($\mathbf{d}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$) for only $2n+1$ equations ($n$ zero-lagrangian equations, $n$ zero-complementarity equations on $\mathbf{d}$, and 1 equation for the active production constraint).

For this reason, one can restrict attention to a subset of the generalized Nash equilibria called normalized Nash equilibria.

**Definition 3.1.** *For a given vector of weights* $\mathbf{w} = (w_1, w_2, \ldots, w_n) > 0$, *a n-tuple* ($\mathbf{d}$, $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$) *is a normalized Nash equilibrium of our production game with respect to* $\mathbf{w}$, *if it satisfies the KKT conditions (3.3) and if there exists a common scalar* $\mu > 0$ *such that:*

$$\mu_i = \mu \; w_i, \qquad \textit{for all} \quad i = 1, \ldots, n$$

Compared to generalized Nash equilibria, a normalized Nash equilibrium has its multipliers $\mu_i$ proportional to each other through the weights $\mathbf{w}$. These weights specify the relative importance of the subsidiaries' contribution to the overall company profit. If $w_i$ is small compared to the other $w_j$'s, then $\mu_i$ will be small compared to the $\mu_j$'s. This means that more leniency is given to subsidiary $i$ in terms of satisfying the energy target. Subsidiary $i$ is allowed to produce more to optimize its profit. On the other hand, a large $\mu_i$ forces subsidiary $i$ to reduce its production by a large amount to satisfy the energy target. Rosen in [113] provides a rigorous proof of the existence and uniqueness of a normalized Nash equiibrium for every given positive weight vector $\mathbf{w}$.

In particular, we will analyze the special case where vector $\mathbf{w}$ is uniform ($w_1 = w_2 = \ldots = w_n$):

**Definition 3.2.** *The normalized Nash equilibrium corresponding to the uniform vector of weights* $\mathbf{w} = (w, w, \ldots, w)$ *is called the* **uniform Nash equilibrium.**

### 3.2.2 Reward-based incentive mechanism

It is common practice in industry (see [96]) that a company establishes an environmental conservation committee to ensure compliance with a committed energy reduction target. This independently funded committee offers a financial reward to each subsidiary for a company-level energy consumption below the target. We assume that the committee first ranks the subsidiaries according to the fraction of the energy reduction burden each of them should bear (or alternatively, according to the importance for the company of each subsidiary's sales level). To this end, the committee uses a weight allocation vector $\mathbf{w} = (w_1, ..., w_n)$ ($w_i > 0$ and $\sum w_i = 1$) that determines the portion of the reward each subsidiary gets. A large $w_i$ results (see below (3.5)) in subsidiary $i$ getting paid a lot per unit of energy saved. Subsidiary $i$ will thus be willing to significantly reduce its energy consumption because its loss of profit due to reduced production is offset by the financial reward it gets for saving energy. Since the $w_i$'s are normalized ($\sum w_i = 1$), a large $w_i$ indicates that the parent company wants to reduce the energy consumption (or the production) of subsidiary $i$ more than the consumption of the other subsidiaries. The parent company might choose a large weight $w_i$ because subsidiary $i$ uses a lot of energy so that the company feels its important to cut that subsidiary's energy consumption. The parent company might also choose the weights $\mathbf{w}$ for strategic reasons. If a certain subsidiary manufactures a product that is more crucial to the company (for example, that subsidiary is the star/luxury brand that reflects the company image), the company might not want to reduce that subsidiary's production too much and will give it a small reward weight $w_i$. Once this allocation vector is chosen, it is viewed as a fixed set of parameters for the rest of the problem.

The committee then needs to decide the reward $\mu$ per unit of energy consumption below the target $C$; each subsidiary will receive a subsidy $\mu\ w_i$ per unit. The committee chooses the level $\mu$ of subsidy and announces it to the subsidiaries. The subsidiaries then choose their production levels to maximize their profit from sales and subsidies. The committee must choose a sufficiently high subsidy level to guarantee that the subsidiaries will satisfy the energy consumption target. At the same time, since the environmental conservation committee is usually also responsible for monitoring other energy reduction activities within the company, it is of its interest to minimize the actual payout of the reward.

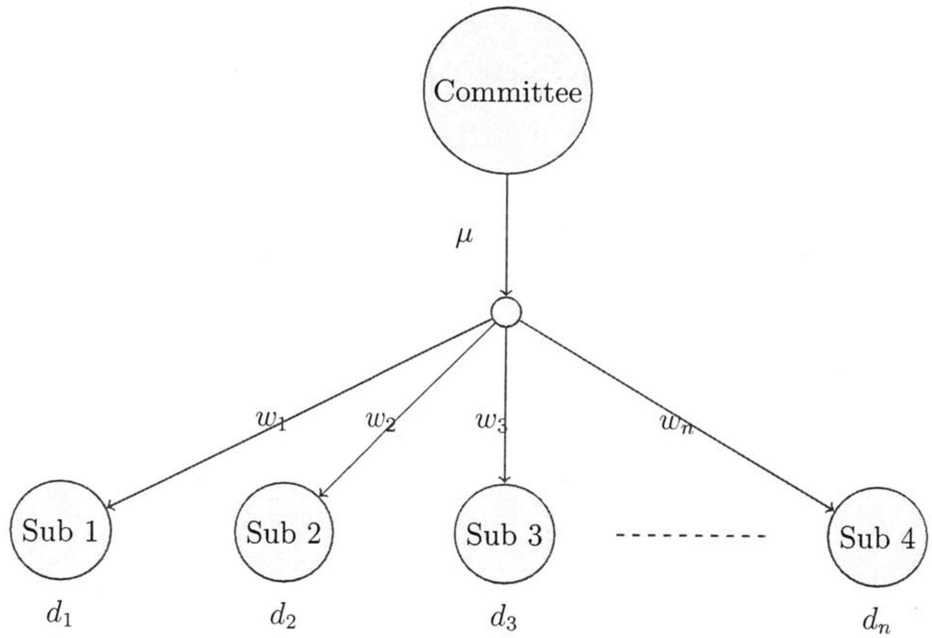


Figure 3-1: Stackelberg subsidy game: $w$'s are set parameters, $\mu$ and $\mathbf{d}$ are the decision variables.

This gives rise to a Stackelberg game between the committee and the subsidiaries. The committee is the leader of the game and anticipates each subsidiary's output as a function of the unit reward $\mu$. Given a predetermined allocation vector $\mathbf{w}$, the committee chooses $\mu$ to minimize actual payments of the reward.

$$\begin{aligned} \mu^* = \underset{\mu}{\operatorname{argmin}} \quad & \mu \left(C - \textstyle\sum_{i=1}^n d_i^*(\mu)\right) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^n d_i^*(\mu) \le C \\ & \mu \ge 0 \end{aligned} \tag{3.4}$$

For subsidiary $i$, the problem is to choose a strategy $d_i^*$ that maximizes its profit (from sales and subsidy) given the reward scheme $(\mathbf{w}, \mu)$ and the strategies of the other subsidiaries $\mathbf{d}^*_{-i}(\mu)^1$:

$$d_i^*(\mu) = \underset{d_i \geq 0}{\text{argmax}} \quad \underbrace{d_i\ (\bar{p}_i - \sum_{j \neq i}^{n} B_{ij} d_j^*(\mu) - B_{ii} d_i)}_{\text{sales}} \ + \ \underbrace{w_i \mu\ (C - \sum_{j \neq i}^{n} d_j^*(\mu) - d_i)^+}_{\text{compliance reward}} \quad (3.5)$$

The committee must decide the reward $\mu$ by solving (3.4) while anticipating that the production of the subsidiaries will solve optimization (3.5). Replacing each subsidiary's optimization (3.5) by it's KKT conditions, the problem of the committee thus becomes an MPEC (maximization problem with equilibrium constraints). MPECs are in general hard to solve, see [?].

Note that one can define a subsidy game for every weight allocation vector $\mathbf{w}$ in the simplex $\{w_i > 0,\ \sum w_i = 1\}$. In particular,

**Definition 3.3.** *The allocation vector $w_1 = w_2 = ... = w_n = \frac{1}{n}$ is called the* **uniform reward allocation scheme.**

### 3.2.3 Equivalence of the models

**Theorem 3.1.** *Every generalized Nash equilibrium of the environmental problem (3.2) can be obtained through a reward based incentive mechanism for an appropriate weight allocation vector $\mathbf{w}$. In particular, the uniform reward scheme $(w_1 = ... = w_n = \frac{1}{n})$ gives rise to the uniform Nash Equilibrium.*

*Proof.* Given a reward-allocation scheme $\mathbf{w}$ (and given the output quantities $d_j^*(\mu)$ of the other subsidiaries $j \neq i$) , the output quantity $d_i^*(\mu)$ of each subsidiary $i$ is a function of the unit reward $\mu$ given by (3.5), i.e.:

$$d_i^*(\mu) = \underset{d_i \geq 0}{\text{argmax}} \quad d_i\ (\bar{p}_i - \sum_{j \neq i}^{n} B_{ij} d_j^*(\mu) - B_{ii} d_i) \ + \ w_i \mu\ (C - \sum_{j \neq i}^{n} d_j^*(\mu) - d_i)^+$$

[1] $\mathbf{d}_{-i}$ to denotes the vector obtained by removing the $i$th element from vector $\mathbf{d}$.

As we will discuss below, the committee will always ensure that the energy target is met, i.e.: $\mathbf{e}^T\mathbf{d}^*(\mu, \mathbf{w}) \leq C$. Without loss of generality, we can thus omit the positive part (function $(.)^+$) in the last term above: $w_i\mu\ (C - \sum_{j\neq i}^n d_j^*(\mu) - d_i)^+$.

The KKT conditions of the above problem are:

$$\left\{\begin{array}{c} \bar{p}_i - \sum_{j\neq i}^n B_{ij}d_j^* - 2B_{ii}d_i^* - w_i\mu + \lambda_i = 0 \\ \lambda_i\ d_i^* = 0 \\ \lambda_i \geq 0,\ d_i^* \geq 0 \\ i = 1, \ldots, n \end{array}\right. \quad \Leftrightarrow \quad \left\{\begin{array}{c} \bar{\mathbf{p}} - (\mathbf{B} + \Gamma)\mathbf{d}^* - \mu\mathbf{w} + \boldsymbol{\lambda} = \mathbf{0} \\ \lambda_i\ d_i^* = 0 \\ \lambda_i \geq 0,\ d_i^* \geq 0 \end{array}\right.$$

where $\mathbf{w}$ and $\boldsymbol{\lambda}$ are n-dimensional vectors.

In particular, the total output level can be expressed as:

$$\begin{aligned} \mathbf{e}^T\mathbf{d}^* &= \mathbf{e}^T(\mathbf{B} + \Gamma)^{-1}(\bar{\mathbf{p}} - \mu\mathbf{w} + \boldsymbol{\lambda}) \\ &= \mathbf{e}^T(\mathbf{B} + \Gamma)^{-1}(\bar{\mathbf{p}} + \boldsymbol{\lambda}) - \mathbf{e}^T(\mathbf{B} + \Gamma)^{-1}\mathbf{w}\mu \end{aligned}$$

Hence $\mathbf{e}^T\mathbf{d}^*$ is a continuous and decreasing function of $\mu$ [2]. Given a fixed consumption target $C$, the committee can always ensure compliance by using a sufficiently large unit reward $\mu$.

The problem of the environmental conservation committee is, as given by (3.4), to minimize subsidy payments while satisfying the energy target:

$$\begin{aligned} \min_{\mu} \quad & \mu\ (C - \mathbf{e}^T\mathbf{d}^*(\mu, \mathbf{w})) \\ \text{s.t.} \quad & \mathbf{e}^T\mathbf{d}^*(\mu, \mathbf{w}) \leq C \\ & \mu \geq 0 \end{aligned}$$

The previous discussion establishes that there exists a large enough $\mu$ that satisfies constraint $\mathbf{e}^T\mathbf{d}^*(\mu, \mathbf{w}) \leq C$. Hence, the subsidy payment is always non-negative for a

[2] This follows since $\mathbf{e}^T(\mathbf{B} + \Gamma)^{-1} = \underbrace{\mathbf{e}^T\mathbf{B}^{-1}}_{>0}\underbrace{\mathbf{B}\ (\mathbf{B} + \Gamma)^{-1}}_{>0} > \mathbf{0}$

feasible $\mu$.

- If $\mathbf{e}^T\mathbf{d}(0, \mathbf{w})) \leq C$, then $\mu = 0$ is optimal since it results in zero payment;

- If $\mathbf{e}^T\mathbf{d}(0, \mathbf{w})) > C$, since $\mathbf{e}^T\mathbf{d}(\mu, \mathbf{w})$ is continuously decreasing in $\mu$, there exists a strictly positive $\mu^*$ satisfying $\mathbf{e}^T\mathbf{d}(\mu^*, \mathbf{w}) = C$. $\mu^*$ is optimal since it also results in zero payments.

Summarizing the above discussion, the solution for the bi-level game is characterized by the following optimality conditions:

$$\left\{\begin{array}{c} \bar{\mathbf{p}} - (\mathbf{B} + \Gamma)\mathbf{d}^* - \mu^*\mathbf{w} + \boldsymbol{\lambda} = \mathbf{0} \\ \mu^* \ (C - \mathbf{e}^T\mathbf{d}^*) = 0 \\ \mathbf{e}^T \ \mathbf{d}^* \leq C \\ \boldsymbol{\lambda} \geq \mathbf{0} \\ \lambda_i \ d_i^* = 0, \ d_i^* \geq 0 \\ \mu^* \geq 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{c} \bar{\mathbf{p}} - (\mathbf{B} + \Gamma)\mathbf{d}^* - \mu^*\mathbf{w} + \boldsymbol{\lambda} = \mathbf{0} \\ \mu^* w_i \ (C - \mathbf{e}^T\mathbf{d}^*) = 0 \\ \mathbf{e}^T \ \mathbf{d}^* \leq C \\ \boldsymbol{\lambda} \geq \mathbf{0} \\ \lambda_i \ d_i^* = 0, \ d_i^* \geq 0 \\ \mu^* \geq 0 \end{array}\right. \tag{3.6}$$

since $w_i > 0$ for $i = 1, 2, ..., n$.

Comparing these equations to the KKT conditions of the generalized Nash equilibrium (3.3) and replacing $\mu_i$ by $\mu \ w_i$, it is clear that the two systems of equations are equivalent. In particular, when $w_1 = \cdots = w_n$, we have $\mu_1 = \cdots = \mu_n$, which by definition, is the uniform Nash-Equilibrium solution. □

### 3.2.4 Company and social objective

If company-wide coordination is possible, the optimal output level is the one that maximizes the sum of all subsidiaries' profits under the committed energy consumption level. We call the corresponding optimization problem the centrally coordinated problem (CP) and use $\mathbf{d}^{CP}$ to denote its solution.

$$
\begin{aligned}
\mathbf{d}^{CP} = \underset{\mathbf{d}}{\operatorname{argmax}} \quad & \mathbf{d}^T(\bar{\mathbf{p}} - \mathbf{B}\ \mathbf{d}) \\
\text{s.t.} \quad & \mathbf{e}^T\ \mathbf{d} \leq C \\
& \mathbf{d} \geq \mathbf{0}
\end{aligned}
\tag{3.7}
$$

From the society's perspective, (that is, considering also the consumers' utility) the optimal production quantity solves the following problem:

$$
\begin{aligned}
\mathbf{d}^{SMAX} = \underset{\mathbf{d}}{\operatorname{argmax}} \quad & CS(\mathbf{d}) + \Pi(\mathbf{d}) = \mathbf{d}^T(\bar{\mathbf{p}} - \tfrac{1}{2}\mathbf{B}\mathbf{d}) \\
\text{s.t.} \quad & \mathbf{e}^T\ \mathbf{d} \leq C \\
& \mathbf{d} \geq \mathbf{0}
\end{aligned}
\tag{3.8}
$$

In the rest of the chapter, we will use this setting as a benchmark to evaluate the performance of the different solutions to the environmental problem (worst-case Nash, uniform Nash, coordinated solution) in terms of social welfare. We refer to the above problem as the SMAX problem and use $\mathbf{d}^{SMAX}$ to denote the corresponding optimal solution.

The price-quantity relationship is modeled through a linear demand function and matrix $\mathbf{B}$ is the inverse of an M-matrix (see [68] for a reference on M-matrices), so that social surplus and company profit are concave functions of the output quantity $\mathbf{d}$. Hence both the SMAX problem and the centrally coordinated problem are maximization problems of a concave function over a simplex constraint. Existence and uniqueness of an optimal solution follows easily.

## 3.3 The Need for Coordination

In this section, we look at the problem from the perspective of the parent company but also from the perspective of society. To highlight the effect of the joint energy constraint, we compare the performance of free competition in the presence of a joint constraint and without it.

As one would expect, free competition between the subsidiaries induces a loss of profit for the company compared to the centrally coordinated scenario. Less intuitive though is the fact that the scale of this loss of profit is much larger in the presence of a joint constraint than without it. The next two theorems show that while the loss of profit due to free competition is bounded in the unconstrained case, it can be arbitrarily bad in the presence of a joint constraint.

**Theorem 3.2.** *In the absence of constraints, the loss of profit resulting from free competition between subsidiaries of a company is bounded by:*

$$\Pi(OP)/\Pi(CP) \geq \max\left\{\frac{2}{2+r\cdot(n-1)};\frac{3}{4+r\cdot(n-1)}\right\}$$

*where $r$ is the diversion ratio defined in (3.1), $n$ is the number of subsidiaries.*

*Proof.* Refer to Appendix B.1. $\square$

In particular, if the company only has two subsidiaries, the loss of profit can be at most 1/3 (using the left side of the bound for a very competitive market $r = 1$, see (3.1)). On the other hand,

**Theorem 3.3.** *In the presence of a joint constraint, free competition between a finite number of subsidiaries can result in an arbitrarily large loss of profit for the company.*

This result can be shown through the following example of duopoly competition:

- $N = 2$
- $\mathbf{B} = \begin{bmatrix} \beta_1 & \alpha \\ \alpha & \beta_2 \end{bmatrix}$, for any $\beta_1 \geq \beta_2 > \alpha \geq 0 \quad \Rightarrow \quad \mathbf{M} = \frac{1}{\beta_1\beta_2-\alpha^2}\begin{bmatrix} \beta_2 & -\alpha \\ -\alpha & \beta_1 \end{bmatrix}$
- $\bar{\mathbf{d}} = \begin{bmatrix} \bar{d}_0 \\ \bar{d}_0 \end{bmatrix}$, for any $\bar{d}_0 > 0 \quad \Rightarrow \quad \bar{p}_1 = (\beta_1+\alpha)\bar{d}_0 \geq (\beta_2+\alpha)\bar{d}_0 = \bar{p}_2$

For a small enough value of the contsrtaint $C$ (namely, $C < \min\left\{\frac{\bar{p}_1-\bar{p}_2}{2(\beta-\alpha)}, \frac{\bar{d}_0}{2}\right\}$), the optimal solution to the SMAX problem and the centrally coordinated problem coincide. It is given by (see Appendix B.1 for the derivation):

$$\mathbf{d}^{SMAX} = \mathbf{d}^{CP} = \begin{bmatrix} c \\ 0 \end{bmatrix}$$

On the other hand, consider a particular equilibrium for the oligopoly problem: $\mathbf{d}^{worst} = \begin{bmatrix} 0 \\ c \end{bmatrix}$. The most inefficient subsidiary produces the entire production. It is an equilibrium since both subsidiaries want to produce more but cannot because of the capacity constraint.

Noting that $r_1 = \alpha/\beta_2$ and $r_2 = \alpha/\beta_1$, we have

$$\frac{r_2}{r_1} = \frac{\beta_2}{\beta_1} \leq \frac{\Pi(\mathbf{d}^{worst})}{\Pi(CP)} = \frac{\beta_2(\bar{d}_0 - C) + \alpha\bar{d}_0}{\beta_1(\bar{d}_0 - C) + \alpha\bar{d}_0} = \frac{r_2}{r_1} \cdot \frac{(r_1+1)\bar{d}_0 - C}{(r_2+1)\bar{d}_0 - C} \tag{3.9}$$

Hence even for two subsidiaries, the loss of profit $\left(1 - \frac{\Pi(\mathbf{d}^{worst})}{\Pi(CP)}\right)$ goes up to 100% as $r_2 \to 0$.

For the rest of the section, we adopt the viewpoint of a regulator, and examine the effect of the joint constraint on social welfare. This includes the firm's profit as well as the consumers' surplus. We compare the unconstrained and the constrained settings.

**Theorem 3.4.** *In the absence of a joint constraint, the loss of social surplus resulting from free competition between subsidiaries of a company is at most 1/4:* $TS(OP) \geq TS(CP) = \frac{3}{4}TS(SMAX)$. *Equality is achieved when subsidiaries are independent.*

*Proof.* Refer to Appendix B.1. □

Without the energy constraint, the maximum loss of social surplus is bounded by a constant (25%) independent of the number of subsidiaries. The bound is achieved when the subsidiaries sell to independent markets, i.e. when they have no influence over the price of their competitiors. In this case, the oligopoly strategy coincides with the centrally coordinated strategy so that the subsidiaries are able to extract the maxmium profit from consumers even in the decentralized setting.

In the presence of a joint constraint, there are multiple oligopoly strategies when the constraint is active. The worst-case loss of social welfare can be arbitrarily bad and the consumers might be better off when the subsidiaries collude to achieve maximum profit than when they compete freely. In this case, colluding induces a more efficient allocation of the scarce resource among subsidiaries which benefits both the firm and society.

**Theorem 3.5.** *In the presence of a joint constraint, free competition between the subsidiaries can result in an arbitrarily large loss of welfare for society.*

This result is shown using the same duopoly example as above. Using the same notations and parameters, we can write:

$$\frac{r_2}{r_1} = \frac{\beta_2}{\beta_1} \leq \frac{TS(\mathbf{d}^{worst})}{TS(SMAX)} = \frac{\beta_2(\bar{d}_0 - 1/2\ C) + \alpha\bar{d}_0}{\beta_1(\bar{d}_0 - 1/2\ C) + \alpha\bar{d}_0} = \frac{r_2}{r_1} \cdot \frac{(r_1+1)\bar{d}_0 - 1/2\ C}{(r_2+1)\bar{d}_0 - 1/2\ C} \tag{3.10}$$

Even with two subsidiaries, the loss of welfare $\left(1 - \frac{TS(\mathbf{d}^{worst})}{TS(SMAX)}\right)$ goes up to 100% as $r_2 \to 0$.

When the capacity constraint is extremely restrictive, both the centrally coordinated solution and the SMAX solution allocate all available capacity to the most efficient subsidiary. In this extreme case, the set of oligopoly solutions on the other hand encompasses all feasible solutions. The maximum loss occurs when all the capacity is allocated to the least efficient subsidiary.

We now turn to companies with specific characteristics and establish bounds on the loss of profit and social welfare under additional assumptions. In particular, we consider companies with fully symmetric subsidiaries and subsidiaries with symmetric price potentials (defined later). We analyze the impact of the joint energy target on the loss of profit and welfare for such companies.

## 3.4 Many Symmetric Subsidiaries

We start the analysis of the performance of deregulated Cournot competition and of the subsidy system by focusing first on the case of fully symmetric subsidiaries. In this section, we consider the situation where the firm's subsidiaries are fully symmetric. By symmetry, we mean that the subsidiaries have the same demand potential and are identical in their abilities to influence each other's demand.

### 3.4.1 Assumptions and closed-form solutions

- All the subsidiaries face the same demand potential: $\bar{\mathbf{d}} = \bar{d}_0\ \mathbf{e}$, for some $\bar{d}_0 > 0$.

- The subsidiaries are identical in their ability to influence each other's price (or demand). With our linear price demand relationship, the assumption translates into matrix $\mathbf{B}$ being uniform. The sensitivity of the subsidiaries' demand to their own price is identical across subsidiaries $B_{11} = B_{22} = ... = B_{nn} = \beta$ for some $\beta > 0$. Similarly, the cross sensitivities, the sensitivity of subsidiaries' demand to competitors' prices, are also uniform $B_{ij} = \alpha$, for all $j \neq i$ for some $0 \leq \alpha < \beta$.

$$\mathbf{B} = \begin{bmatrix} \beta & \alpha & \dots & \alpha \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ \alpha & \dots & \alpha & \beta \end{bmatrix} = \mathbf{M}^{-1} = \begin{bmatrix} M & -m & \dots & -m \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ -m & \dots & -m & M \end{bmatrix}^{-1}$$

  for some $M > (n-1)\ m \geq 0$.

- $\bar{\mathbf{p}} = \mathbf{B}\bar{\mathbf{d}} = \bar{p}_0\mathbf{e}$, where $\bar{p}_0 = (\beta + (n-1)\ \alpha)\ \bar{d}_0$;

- the diversion ratio becomes: $r = \frac{(n-1)\ m}{M} = \frac{(n-1)\ \alpha}{(n-2)\ \alpha\ +\ \beta}$

The closed-form solutions are summarized as follows (refer to Appendix B for a proof ):

$$\begin{array}{rl} d_i^{CP} = & \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{2(\beta+(n-1)\alpha)}\right\} \\ d_i^{SMAX} = & \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{\beta+(n-1)\alpha}\right\} \\ d_i^{UNE} = & \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{2\beta+(n-1)\alpha}\right\} \end{array} \quad \forall i = 1, 2, \cdots, n.$$

$$\mathbf{d}^{OP} = \begin{cases} \frac{\bar{p}_0}{2\beta+(n-1)\alpha}\mathbf{e} & \text{, when } \frac{\bar{p}_0}{2\beta+(n-1)\alpha} \leq \frac{C}{n}; \\ \left\{\mathbf{d} \mid 0 \leq d_i \leq \frac{\bar{p}_0-\alpha C}{2\beta-\alpha}, \sum d_i = C\right\} & \text{, otherwise.} \end{cases}$$

When capacity is restrictive, the centrally coordinated solution, SMAX solution and uniform Nash equilibrium coincide: $\mathbf{d} = \frac{C}{n}\mathbf{e}$. When $C \leq \frac{\bar{p}_0}{2\beta}$, the set of oligopoly equilibria, on the other hand, encompasses all possible production repartition between the subsidiaries in which total production reaches capacity.

### 3.4.2 Loss of profit and welfare

We first look at the loss of profit resulting from free competition.

**Theorem 3.6.** *For symmetric subsidiaries facing a single joint capacity constraint, the fraction of profit achieved under free competition compared to the maximum company profit is at least:*

- *When the capacity constraint* **is not active for both** *the centrally coordinated problem and the oligopoly problem (i.e., when $C \geq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}$ ), then:*

$$\frac{\Pi(OP)}{\Pi(CP)} = \Phi_1(r,n) \geq \frac{4n}{(n+1)^2} \tag{3.11}$$

- *When the capacity constraint* **is active for both** *problems (i.e., when $C \leq \frac{\bar{p}_0}{2\alpha+\frac{2(\beta-\alpha)}{n}}$ ), then:*

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \Phi_2(r,n) > \frac{1}{2} \tag{3.12}$$

- *When the constraint* **is active for the oligopoly** *problem but* **inactive for the centrally coordinated** *problem, the profit ratio lies between the unconstrained bound (3.11) and the constrained bound (3.12).*

Figure 3-2 and 3-3 illustrate the bounds provided by $\Phi_1$ and $\Phi_2$.

*Proof.* The derivation of the bounds as well as the expressions of $\Phi_1(r, n)$ and $\Phi_2(r, n)$ can be found in Appendix B.2. □

When the capacity constraint is inactive, the loss of profit is bounded and (3.11) describes how this loss of profit varies with the number of subsidiaries $n$. When the constraint is tight on the other hand, inequality (3.9) shows that for general markets (with no specific characteristics) the loss of firm profit resulting from free competition between the subsidiaries can be arbitrarily large even with only two subsidiaries. In contrast, (3.12) establishes that when the constraint is tight, if the subsidiaries are fully symmetric, the profit loss cannot exceed 1/2 no matter the number of subsidiaries, the intensity of competition or the value of the active constraint. If the subsidiaries are similar enough, the company needs to compare the cost of implementing a subsidy system to coordinate its subsidiaries with the potential loss of half their profit.

Figure 3-2 plots the profit loss $(1 - \Phi_1(r, n))$ for the worst oligopoly equilibrium in the unconstrained case as a function of the number of subsidiaries $n$ for various valus of $r$. The loss of profit increases with the number of subsidiaries and the intensity of competition. As competition intensifies, the profit loss becomes more sensitive to the increase in $n$. In case of fierce competition ($r$ close to 1), the profit loss is 12% with two subsidiaries and 67% with 10 subsidiaries. These trends are intuitive as the more subsidiaries there are and the more intense competition is, the bigger the unaccounted for externality each subsidiary inflicts on the others. The function $1 - \Phi_1$ depicted in Figure 3-2 is the exact value of the worst-case profit loss with unlimited capacity and the ratio $\frac{4n}{(n+1)^2}$ is a tight bound reached under fierce competition ($r \to 1$).

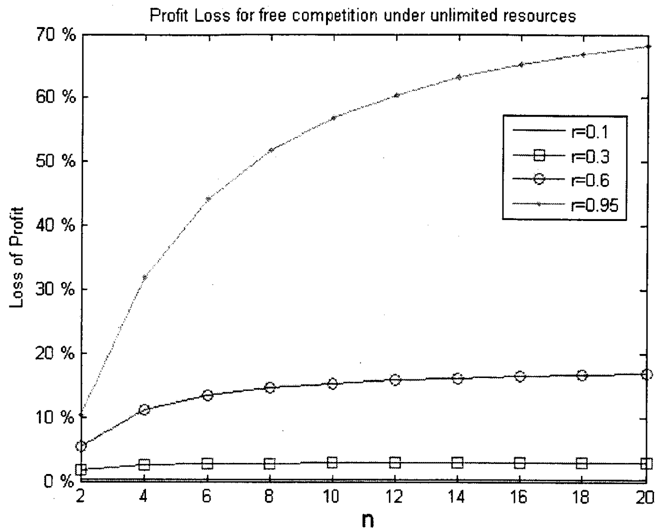


Figure 3-2: Profit loss $(1 - \Phi_1(r, n))$ for free competition under unlimited resources

Figure 3-3 plots the profit loss $(1 - \Phi_2(r, n))$ for the worst oligopoly equilibrium in the tightly constrained case (i.e., when $C = \frac{\bar{p}_0}{2\beta}$ ) as a function of the number of subsidiaries $n$ for various values of $r$. The loss of profit again increases with the number of subsidiaries but it now decreases with the intensity of competition. We will explain this seemingly counterintuitive trend at the end of the section when analyzing the actual worst case production vector of the subsidiaries.

The bound $1 - \Phi_2(r, n)$ is tight for a discrete set of capacity values. To visualize this, we compare in Figure 3-4, $1 - \Phi_2$ and the actual worst-case profit loss as functions of capacity when $n = 10$ and $r = 0.1$. The bound is tight at every value of the capacity where the two curves meet. As expected, the bound $\Phi_2$ stays below 50%; that is the loss of profit never exceeds a half. The profit loss is highly non-monotonic as a function of capacity. Although for larger values of the capacity, the profit loss globally decreases at the same rate as the bound, within each tightness interval ($C \in \left( \frac{\bar{p}_0}{\alpha + \frac{2\beta - \alpha}{k}}, \frac{\bar{p}_0}{\alpha + \frac{2\beta - \alpha}{k+1}} \right]$) the profit loss remains non-monotonic.

In order to understand the effect of a joint constraint, Figure 3-5 compares the worst case loss of profit in the constrained case with the loss of profit in the unconstrained case, The dotted line represents the constrained loss of profit under the most

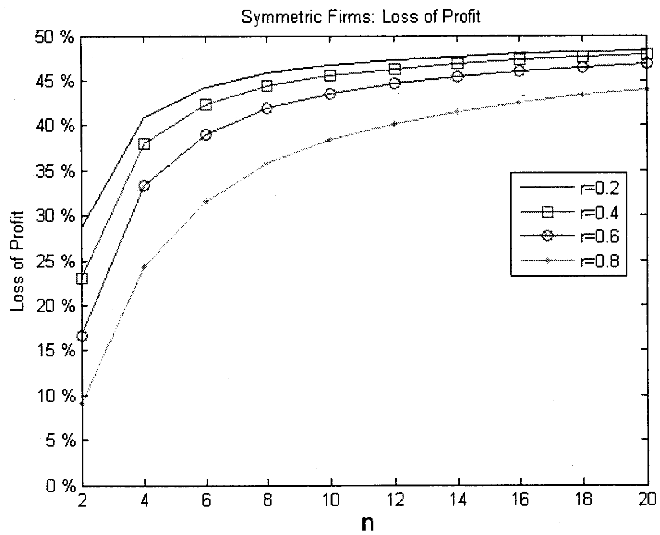


Figure 3-3: Maximum profit loss $(1 - \Phi_2(r, n))$ for free competition under limited resources

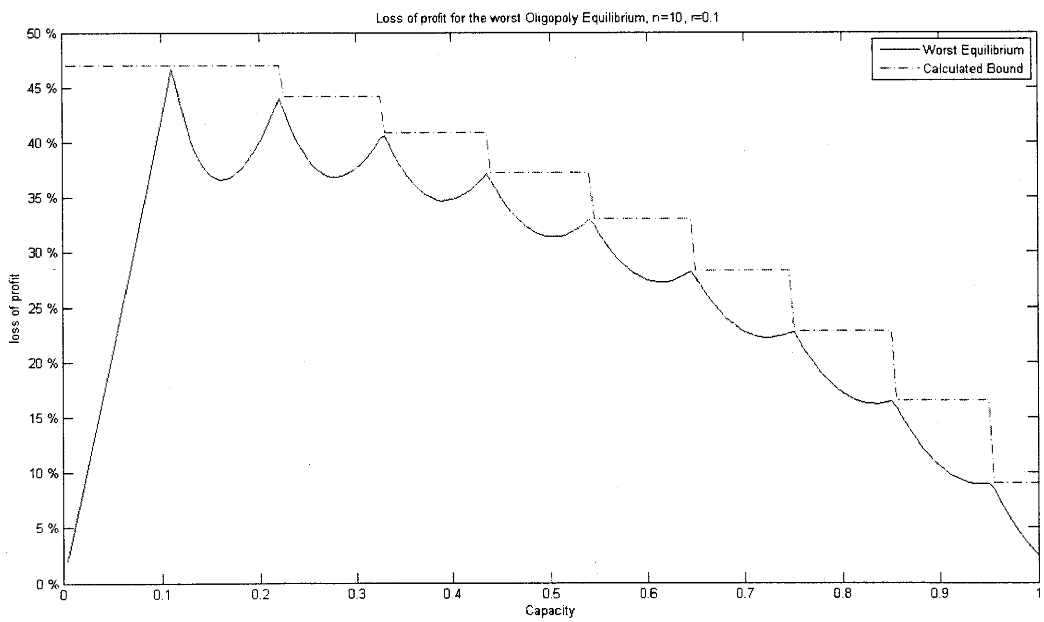


Figure 3-4: Profit loss for the worst equilibrium

damaging value of capacity and the squared curve represents the unconstrained loss of profit. These losses are plotted as functions of the diversion ratio $r$ for various number of competing subsidiaries $n$. The general insight is that the profit loss is generally much larger in the constrained case. Coordinating the subsidiaries through a subsidy system is even more valuable in the presence of an energy consumption constraint. In the presence of intense competition ($r \to 1$) however, this relationship is reversed: the profit loss is larger in the unconstrained case. As will become clearer at the end of the section, while competition between subsidiaries causes mainly negative externalities in the unconstrained case, it induces some positive externalities in the presence of a joint constraint.

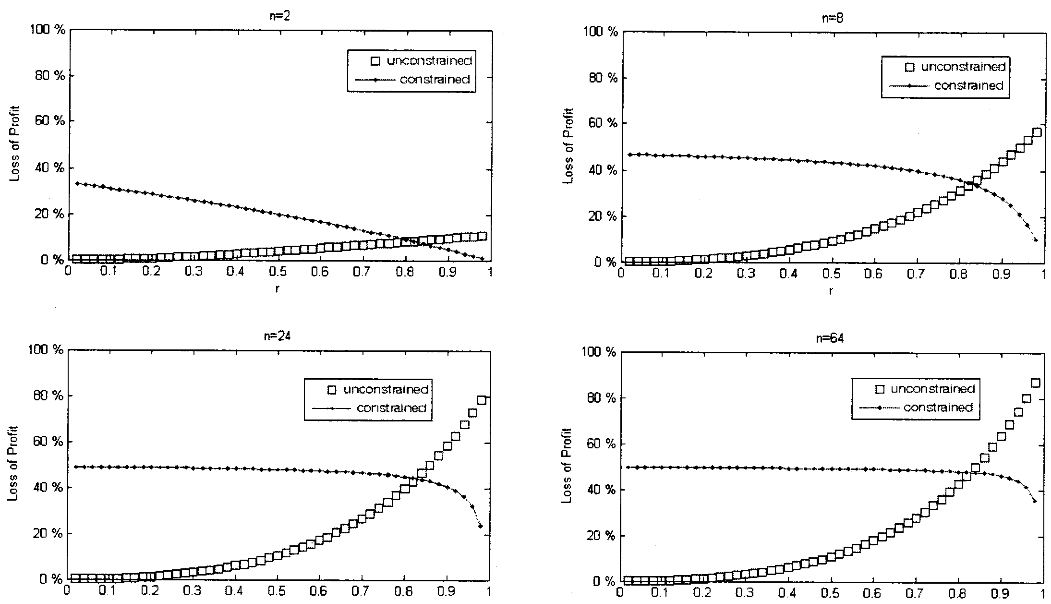


Figure 3-5: Effect of the joint constraint on the loss of profit

Next, we evaluate the benefit of free competition for society by measuring social welfare. We benchmark the total surplus under free competition against the maximum achievable total surplus.

**Theorem 3.7.** *For symmetric subsidiaries facing a single joint capacity constraint, the fraction of social welfare achieved under free competition compared to the maximum achievable welfare is at least 75%:*

- *When the capacity constraint* **is not active for both** *the oligopoly problem and* *the SMAX problem (i.e., when* $C \geq \frac{\bar{p}_0}{\alpha + \frac{\beta - \alpha}{n}}$*), then:*

$$\frac{TS(OP)}{TS(SMAX)} = \Phi_3(r, n) \geq \frac{3}{4}; \tag{3.13}$$

- *When the constraint* **is active for both** *problems (i.e., when* $C \leq \frac{\bar{p}_0}{\alpha + \frac{2\beta - \alpha}{n}}$*), then:*

$$\frac{TS(OP)}{TS(SMAX)} \geq \Phi_4(r, n) > \frac{3}{4} \tag{3.14}$$

- *When the constraint* **is active for the SMAX** *problem but* **not for the oligopoly** *problem the social welfare ratio lies above the unconstrained bound (3.13).*

Figure 3-6 illustrates the bound provided by $\Phi_4$.

*Proof.* The derivation of the bounds as well as the expressions $\Phi_3(r, n)$ and $\Phi_4(r, n)$ can be found in Appendix B.2. □

Whereas without any further assumption on the market characteristics, the loss of social welfare resulting from free competition between the subsidiaries can be arbitrarily large (see (3.10)), Theorem 3.7 shows that this loss is bounded by 3/4 if the subsidiaries are symmetric regardless of the number of subsidiaries or the intensity of competition. From the point of view of a regulator (or society), letting the subsidiaries compete is a good option.

Figure 3-6 plots the maximum loss of social welfare $(1 - \Phi_4(r, n))$ in the extremely constrained case (i.e., when $C = \frac{\bar{p}_0}{2\beta}$) as a function of the number of subsidiaries $n$ for different diversion ratios $r$. The loss of social welfare increases with the number of subsidiaries but decreases with the intensity of competition. It follows similar trends as the constrained loss of profit; we explain these trends at the end of the section.

Function $\Phi_3$ is the exact value of the worst-case social welfare loss with unlimited capacity and the ratio $\frac{3}{4}$ is a tight bound reached when the subsidiaries are independent (i.e., $r \to 0$). The constrained loss of social welfare bound $\Phi_4(r, n)$ is tight on the

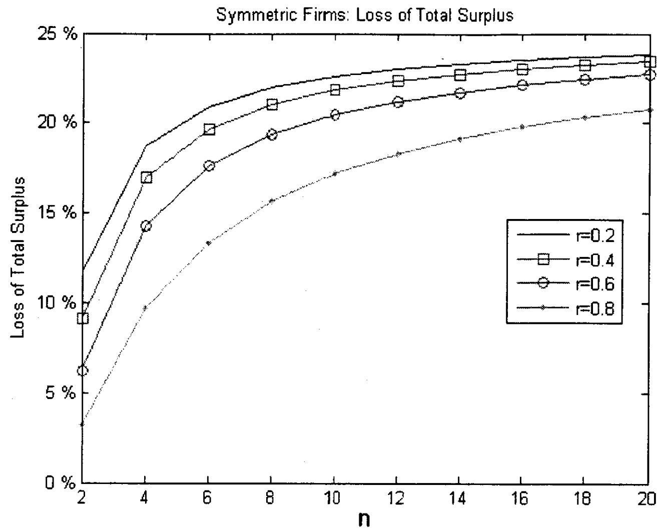


Figure 3-6: Maximum loss of social welfare $(1 - \Phi_4(r, n))$ under free competition with limited resources

same discrete set of capacity values as the constrained loss of profit bound $\Phi_2(r, n)$. The maximum loss of 25% is achieved when the market consists of numerous independent subsidiaries (i.e., $r \to 0, n \to \infty$).

We highlight the effect of the joint constraint by comparing in Figure 3-7 the worst loss of social welfare in the constrained case, with the loss of social welfare in the unconstrained case. The dotted line represents the constrained loss of social welfare under the most damaging value of capacity and the squared curve represents the unconstrained loss of social welfare. These losses are plotted as functions of the diversion ratio $r$ for various number of competing subsidiaries $n$. While the worst loss of social welfare in the constrained case is larger than the unconstrained loss when the number of competing subsidiaries is large, competition under limited resources is more efficient when the firm only has a few subsidiaries regardless of the intensity of competition.

In summary, when the capacity constraint is active, free competition induces a loss of profit for the company of no more than 1/2 and a loss of social welfare of no more than 1/4. How does the simple uniform reward scheme perform in this setting?

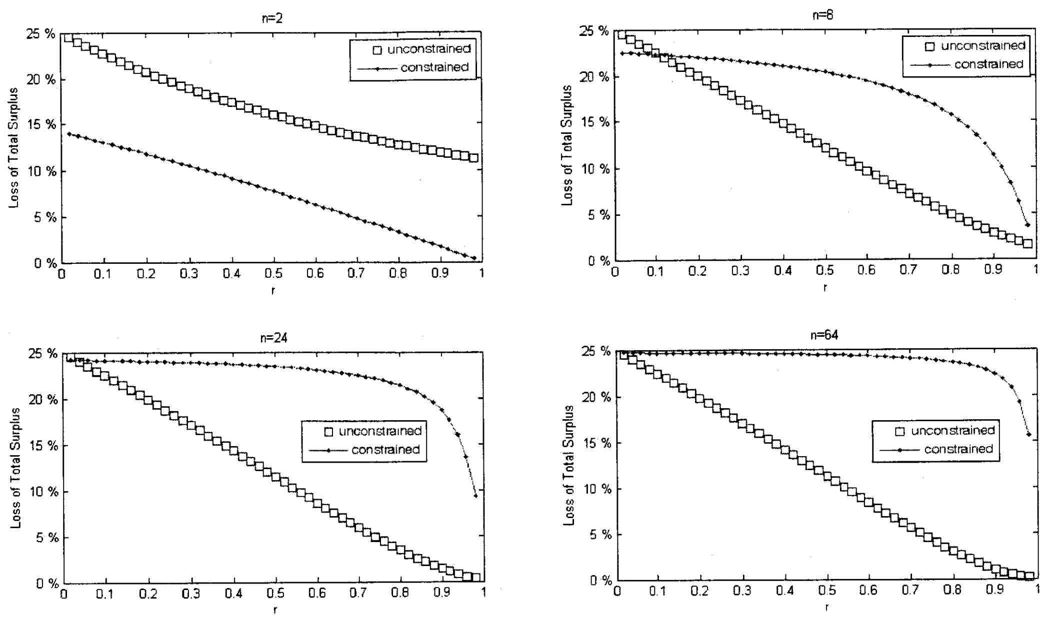


Figure 3-7: Effect of the joint constraint on the loss of social welfare

Does it improve company profit? Does it hurt consumers?

**Theorem 3.8.** *When the capacity constraint* **is active for the SMAX**, *the centrally coordinated and the oligopoly problems, the uniform Nash equilibrium is the best oligopoly equilibrium. Moreover, it achieves maximum possible profit and social welfare:*

$$\frac{TS(UNE)}{TS(SMAX)} = 1 \qquad \textit{and} \qquad \frac{\Pi(UNE)}{\Pi(CP)} = 1$$

*Proof.* The theorem follows since $\mathbf{d}^{SMAX} = \mathbf{d}^{CP} = \mathbf{d}^{UNE} = \begin{bmatrix} \frac{C}{n} \\ \vdots \\ \frac{C}{n} \end{bmatrix}$ when the constraint is active. $\square$

Not surprisingly, if the subsidiaries are similar enough, enticing them to split production evenly through the uniform reward scheme is the optimal strategy for the firm. This behavior is also optimal for society; the regulator (who cares about social welfare) does not need to worry about consumers getting hurt and is happy to see the company coordinating its subsidiaries.

### 3.4.3 Underlying production strategies

In the absence of constraints, the loss of social welfare under free competition is caused by underproduction compared to the SMAX solution and the loss of profit is caused by overproduction compared to the centrally coordinated solution: $d_i^{SMAX} > d_i^{OP} \geq d_i^{CP}$. In these three solutions, the subsidiaries split production evenly since the market is symmetric. When the subsidiaries sell to independent markets ($r = 0$), free competition achieves optimal profit; the production quantity that is individually optimal for each subsidiary is optimal for the firm as a whole. When the subsidiaries are not independent but compete in the same market ($r \neq 0$), however, the sales of one subsidiary negatively impact the sales of the others. In this case, the output quantity individually chosen by a subsidiary is "too much" from the point of view of the firm because the subsidiary does not account for the reduction in demand it causes to the other subsidiaries. The loss of profit, in this competitive setting, is due to the subsidiaries overproducing.

As the capacity constraint (energy target) becomes restrictive, these different solutions all produce the same total quantity: the parent company as a whole produces at capacity. Under free competition however, there is no longer a unique equilibrium strategy. Among the set of equilibria, the subsidiaries produce evenly only under the uniform Nash equilibrium. Since the total production quantity is constant when the constraint is active, the profit of the parent company depends on the average market price:

$$\Pi(\mathbf{d}) = \sum_i^n d_i p_i(\mathbf{d}) = C\ \mu_p\ ,$$

where $\mu_p = \frac{\sum_i^n d_i\ p_i(\mathbf{d})}{\sum_{i=1}^n d_i}$ is the average market price. Since the subsidiaries are symmetric and price is a concave decreasing function of quantity, the average market price is maximized when each subsidiary sells an equal portion of the total quantity: $\frac{C}{n}$. The average market price is minimized when all the products are sold through a single subsidiary. As a result, the uniform Nash equilibrium is the best equilibrium in terms of profit; it actually coincides with the centrally coordinated solution. The worst equilibrium is the one with the least number of producing subsidiaries. The consumer on the other hand, would prefer to purchase from fewer subsidiaries with a lower price. Hence the worst equilibrium in terms of profit is actually the best in terms of consumer surplus. As competition increases ($r$ increases), the market price decreases and it decreases faster when all the subsidiaries produce evenly than when only a few subsidiaries split production. This causes the optimal profit (under uniform production) to decrease faster than the worst-case profit. The loss of profit $\left(= 1 - \frac{\Pi(OP)}{\Pi(CP)}\right)$ decreases with $r$ as a result. The SMAX problem is a trade off between the parent company's preference for higher market price and the consumers' preference for lower market price. The socially optimal solution coincides with the centrally coordinated solution. It indicates that the loss of firm profit resulting from lower market price is far more significant than the surplus gained by consumers. The loss of social welfare follows the same trends as the loss of profit; it decreases with $r$ for the reasons discussed above.

## 3.5 Many Subsidiaries with Symmetric Price Potentials

In this section, we partially relax the uniformity assumption imposed in the previous section to consider subsidiaries with only uniform price potentials: $\bar{\mathbf{p}} = \mathbf{B}\,\bar{\mathbf{d}} = \mathbf{e}\,\bar{p}_0$, for some $\bar{p}_0 > 0$. However matrix $\mathbf{B}$ (where $\mathbf{p}(\mathbf{d}) = \widetilde{\mathbf{p}} - \mathbf{B}\cdot\mathbf{d}$) is no longer uniform, i.e. subsidiaries no longer have the same price sensitivities. It is still the inverse of a symmetric M-matrix but its diagonal coefficients are free to differ from each other and so do its off-diagonal coefficients. This means that consumers derive the same utility from buying the first unit of any of these products but that their marginal utility decreases faster for further units of certain products than of others.

### 3.5.1 Closed-form solutions

The closed-form solutions for the SMAX, centrally coordinated and uniform Nash equilibrium are given by:

**Lemma 3.1.**

$$\mathbf{d}^{SMAX} = \min\left\{\bar{p}_0\mathbf{B}^{-1}\boldsymbol{e}, \frac{C}{\boldsymbol{e}^T\mathbf{B}^{-1}\boldsymbol{e}}\mathbf{B}^{-1}\boldsymbol{e}\right\}$$

$$\mathbf{d}^{CP} = \min\left\{\frac{1}{2}\bar{p}_0\mathbf{B}^{-1}\boldsymbol{e}, \frac{C}{\boldsymbol{e}^T\mathbf{B}^{-1}\boldsymbol{e}}\mathbf{B}^{-1}\boldsymbol{e}\right\}$$

$$\mathbf{d}^{UNE} = min\left\{\bar{p}_0(\mathbf{B}+\Gamma)^{-1}\boldsymbol{e}, \frac{C}{\boldsymbol{e}^T(\mathbf{B}+\Gamma)^{-1}\boldsymbol{e}}(\mathbf{B}+\Gamma)^{-1}\boldsymbol{e}\right\}$$

*Proof.* Refer to Appendix B.3 for a proof. □

### 3.5.2 Loss of profit and welfare

We examine the loss of profit and welfare under the relaxed assumption of symmetric price potentials but asymmetric price sensitivities.

**Theorem 3.9.** *When the subsidiaries have symmetric price potentials and the capacity constraint* **is active for all problems**, *the losses of profit and welfare for the worst oligopoly equilibrium are characterized by:*

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{1}{2} + \frac{1}{2}\frac{1}{(2B_{MM}(\boldsymbol{e}^T\Gamma^{-1}\boldsymbol{e}) - 1)} > \frac{1}{2} \tag{3.15}$$

$$\frac{TS(OP)}{TS(SMAX)} \geq \frac{3}{4} + \frac{3}{4}\frac{1}{(4B_{MM}\boldsymbol{e}^T\Gamma^{-1}\boldsymbol{e} - 1)} > \frac{3}{4} \tag{3.16}$$

*where* $B_{MM} = \max_i\{B_{ii}\}$.

With uniform price potentials only (non uniform sensitivities), the worst case loss of profit is still 1/2 and the worst case loss of welfare is still 1/4 as in the previous section (see also (3.12) and (3.14)). Hence, dropping the assumption of symmetric price sensitivities does not hurt the worst case losses. Tightness for both bounds is achieved when capacity is tight, i.e., when $C \leq \frac{\bar{p}_0}{2B_{MM}}$. Regardless of the number of firm subsidiaries $n$ or the diversion ratio $r$, the maximum loss of profit and welfare occur when the self sensitivity of one subsidiary $B_{MM}$ is significantly larger than the self sensitivity of the others. This translates into having one diagonal element of matrix $\mathbf{B}$ (the $M$-th diagonal element) much larger than the other diagonal elements of $\mathbf{B}$. It represents a market where the consumers of subsidiary $M$ are very sensitive to price (to product $M$'s price), while the consumers of other subsidiaries aren't as sensitive to the price set by their subsidiary.

In this context, how does the uniform reward allocation scheme perform? We first study the value for the firm of implementing this reward scheme and then evaluate its impact on social welfare.

**Theorem 3.10.** *When the joint constraint* **is active** *and the price potentials are symmetric across subsidiaries, the loss of company profit at the uniform Nash equilibrium compared with the centrally coordinated solution is no more than 1/3:*

$$\frac{\Pi(UNE)}{\Pi(CP)} \geq 2 - 2\delta + \frac{3}{4}\delta^2 \geq \frac{2}{3}, \tag{3.17}$$

*where $\delta$ satisfies: $0 < 2 - r \leq \delta \leq 2$. Tightness of the bound is achieved when $r = 0$.*

Without the assumption of symmetric price sensitivity, the uniform Nash equilibrium is no longer an optimal solution for the firm (in contrast with Theorem 3.8). However, under the uniform reward scheme the worst loss of profit is $1/3$ whereas under free competition between the subsidiaries, this loss can go up to $1/2$. The firm must hence decide if this improvement in profit is worth the cost of implementing the reward system.

Figure 3-8 compares the derived lower bound of the ratio $\frac{\Pi(UNE)}{\Pi(CP)}$ with simulations. The simulations represent the smallest ratio $\frac{\Pi(UNE)}{\Pi(CP)}$ out of 1000 randomly generated examples for each $n$ and $r$. The loss of profit under the uniform Nash equilibrium is independent of the number of subsidiaries. The loss worsens as competition intensifies ($r$ increases) and it is a nondecreasing function of capacity. The maximum loss occurs when the unconstrained centrally coordinated solution exactly reaches capacity.

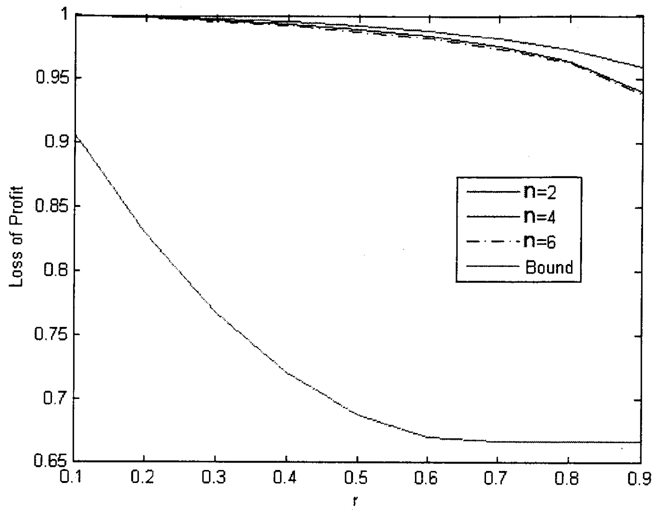


Figure 3-8: Loss of profit for the uniform Nash equilibrium

We now discuss the regulator's perspective by analyzing social welfare.

**Theorem 3.11.** *When the joint constraint* **is active** *and the price potentials are symmetric across subsidiaries, the loss of social welfare at the uniform Nash equilibrium compared with the centrally coordinated solution is no more than 1/4:*

$$\frac{TS(UNE)}{TS(SMAX)} \geq \max\left\{\frac{2}{3} + \frac{2}{3\,(2 + r(n-1))}, \frac{3}{8}\frac{(2-r)^2}{\frac{3}{2} - r}\right\} \geq \frac{3}{4}. \tag{3.18}$$

*The first bound dominates when $n$ is small while the second bound dominates when $n$ is large.*

Without the assumption of symmetric price sensitivity, the uniform Nash equilibrium solution is no longer socially optimal (in contrast with Theorem 3.8). However, compared to free competition, the uniform reward scheme improves the total profit of the firm without deteriorating the social welfare performance. The regulator has no basis to argue against the implementation of the reward system since it does not hurt society.

Figures 3-9, 3-10 and 3-11 compare the derived lower bound of the ratio $\frac{TS(UNE)}{TS(SMAX)}$

with the actual simulated ratios. The loss of social welfare decreases with the number of subsidiaries $n$ and the intensity of competition $r$. Although the bound is only tight when $r = 0$, the simulated curves follow the same trends as the bounds. The loss of social welfare is also a nondecreasing function of capacity; the maximum loss occurs in the totally unconstrained case.

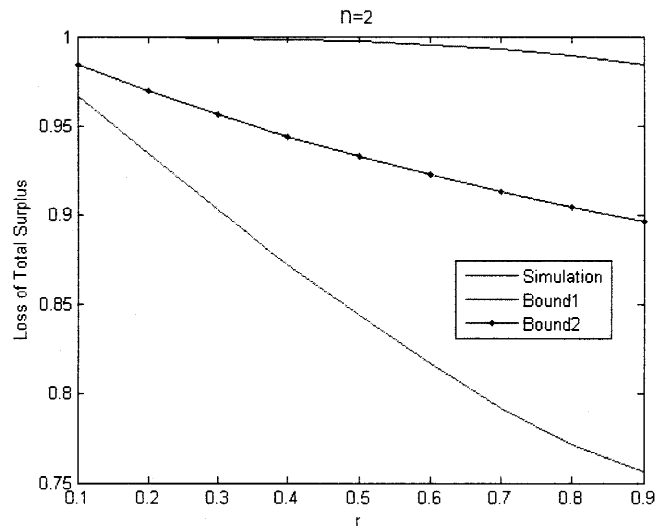


Figure 3-9: Loss of social welfare for the uniform Nash equilibrium: n=2

### 3.5.3 Underlying production strategies

Under the assumption of symmetric price potentials but asymmetric price sensitivities, the SMAX solution still coincides with the centrally coordinated solution when capacity is restrictive. However, the uniform Nash equilibrium is no longer the best equilibrium except for 2 special cases:

1. when the subsidiaries sell to independent markets, so that $\mathbf{d}^{CP} = \mathbf{d}^{UNE}$, the production allocation under uniform reward coincides with the centralized allocation;

2. when the price influence (matrix $\mathbf{B}$) is uniform, so that the market is fully symmetric as in the previous section.

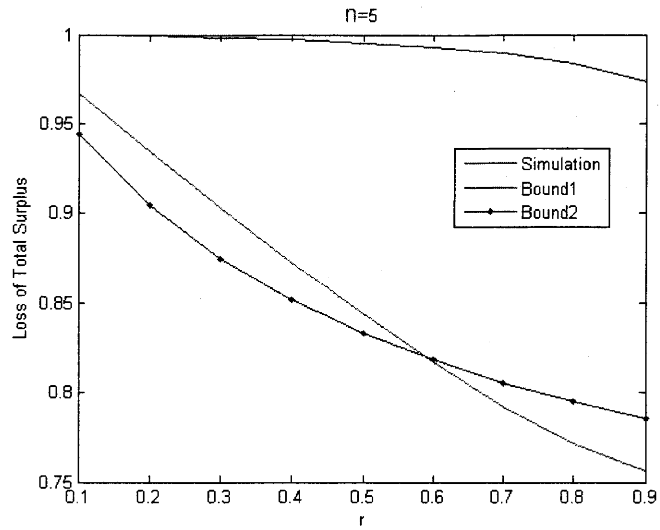


Figure 3-10: Loss of social welfare for the uniform Nash equilibrium: n=5

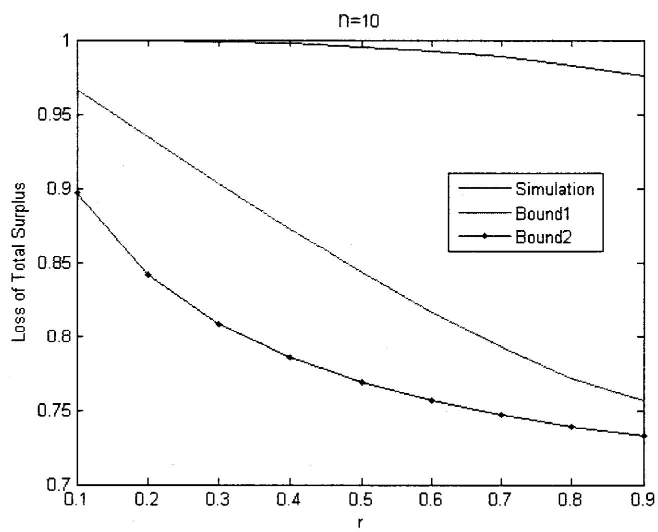


Figure 3-11: Loss of social welfare for the uniform Nash equilibrium: n=10

Similar to the fully symmetric market, while both the centrally coordinated and the SMAX problem try to maximize the average market price, the worst oligopoly equilibrium is the one that results in the minimum market price. With symmetric price potentials, the selling price is given by

$$p_i(\mathbf{d}) = \bar{p}_0 - \sum_{j=1}^{n} B_{ij} d_j$$

so that the smallest possible market price to sell quantity $C$ is $p_{min} = \bar{p}_0 - B_{MM}\ C$, where $B_{MM} = \max\{B_{ii}\}$. In this case, all products are sold through subsidiary M. Subsidiary M is the least efficient one in the market in the sense that if all the subsidiaries operate in their own niche market, the market price of subsidiary M's product is lower than the market price of other subsidiaries' product for the same output level due to possibly inferior quality or other factors that decrease customers' perceived value for the product.

## 3.6 Extension to General Price Potentials

We now drop the assumption of symmetric price potentials. Under this general case, we extend our previous bounds on the loss of company profit and social welfare by introducing an additional measure $\theta = \bar{p}_{min}/\bar{p}_{max}$ of the asymmetry of the price potentials. We also establish some positive results on the centrally coordinated approach.

Section 3.3 showed that in the presence of an active joint constraint the loss of company profit and social welfare resulting from competition between subsidiaries can be arbitrarily large even for a fixed number of subsidiaries. When the subsidiaries have symmetric price potentials, on the other hand, Section 3.5 demonstrated that the loss of profit and social welfare, resulting from free competition, are bounded by 1/2 and 3/4 respectively. We define $\theta = \bar{p}_{min}/\bar{p}_{max}$ to measure the asymmetry between the price potentials of the subsidiaries. $\theta \to 0$ is the most asymmetric case (either $\bar{p}_{min} = 0$ or $\bar{p}_{max}$ is much larger than $\bar{p}_{min}$); $\theta \to 1$ represents the case of symmetric price potentials (as in Section 3.5). Using $\theta$, we generalize Theorem 3.9 and bound the loss of profit and social welfare in the general case.

**Theorem 3.12.** *When the capacity constraint is active, the losses of profit and welfare* *for the worst oligopoly equilibrium are characterized by:*

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{\bar{p}_{min} - B_{MM}C}{\bar{p}_{max}} > \theta - \frac{1}{2} \qquad \textit{for} \quad \bar{p}_{max} \leq 2\bar{p}_{min} \tag{3.19}$$

$$\frac{TS(OP)}{TS(SMAX)} \geq \frac{\bar{p}_{min} - 1/2B_{MM}C}{\bar{p}_{max}} > \theta - \frac{1}{4} \qquad \textit{for} \quad \bar{p}_{max} \leq 4\bar{p}_{min} \tag{3.20}$$

*where* $B_{MM} = \max\{B_{ii}\}$.

*Proof.* Refer to Appendix B.4 for a proof. □

These bounds coincide with results (3.15) and (3.16) of the previous section, when the price potentials are symmetric (i.e. $\theta = 1$). These bounds are also tight: the following duopoly example reaches the bound.

Take two subsidiaries with price potentials $\bar{p}_{min}$ and $\bar{p}_{max}$ respectively, and with a demand sensitivity matrix $\mathbf{B} = \begin{pmatrix} b & 0 \\ 0 & \epsilon \end{pmatrix}$. We assume $\bar{p}_{max} > \bar{p}_{min} \geq 2\ b\ C$ and $\epsilon \ll 1$. Attributing all the production to subsidiary 1 is the worst oligopoly. The centralized and the socially optimal solution, on the other hand, is to have subsidiary 2 produce the entire production quantity $C$. The profit ratio is:

$$\begin{aligned} \frac{\Pi(OP)}{\Pi(CP)} &= \frac{C(\bar{p}_{min} - b\ C)}{C(\bar{p}_{max} - \epsilon\ C)} \quad \underset{\epsilon \to 0}{\rightarrow} \quad \frac{\bar{p}_{min} - b\ C}{\bar{p}_{max}} \\ &> \theta - \frac{1}{2} \qquad\qquad (\text{since } \bar{p}_{max} > 2\ b\ C) \end{aligned}$$

Similarly, the total surplus ratio is:

$$\begin{aligned} \frac{TS(OP)}{TS(SMAX)} &= \frac{C(\bar{p}_{min} - 1/2\ b\ C)}{C(\bar{p}_{max} - 1/2\ \epsilon\ C)} \quad \underset{\epsilon \to 0}{\rightarrow} \quad \frac{\bar{p}_{min} - 1/2\ b\ C}{\bar{p}_{max}} \\ &> \theta - \frac{1}{4} \qquad\qquad (\text{since } \bar{p}_{max} > 2\ b\ C) \end{aligned}$$

In light of these theorems, when the subsidiaries have very asymmetric price potentials ($\bar{p}_{min}/\bar{p}_{max}$ small), free competition performs poorly in terms of both profit and welfare. Hence, it might seem like the only sensible solution for the company as well as for society is to centrally coordinate the subsidiaries. However, such coordination is not always possible (due to regulations or practical considerations) and when it is, it is often extremely costly for the parent company. We demonstrate next that the subsidy system can solve this problem by leading the subsidiaries to adopt provably good solutions in a decentralized manner.

As noted in Section 3.2, every oligopoly solution can be attained with a subsidy system by giving appropriate relative weights to the subsidiaries. We now show that the centrally coordinated solution belongs to the set of oligopoly equilibria. It is hence implementable through a subsidy system.

**Theorem 3.13.** *When the constraint* **is active** *for the centrally coordinated solution ($\mathbf{e}^T\mathbf{d}^{CP} = C$), $\mathbf{d}^{CP}$ belongs to the set of equilibrium strategies of the oligopoly problem.*

*Proof.* To prove this theorem, we first show that when the constraint is active for the centrally coordinated problem, it must be active for the oligopoly problem as well.

This follows since in the unconstrained case, $\mathbf{d}^{OP} = (\mathbf{B}+\Gamma)^{-1}\mathbf{B}\ \bar{\mathbf{d}}$ and $\mathbf{d}^{CP} = \frac{1}{2}\bar{\mathbf{d}}$,

$$\begin{aligned}
\mathbf{e}^T\mathbf{d}^{OP} &= \mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}\ \bar{\mathbf{d}} \\
&= \frac{1}{2}\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}(2B)\bar{\mathbf{d}} \\
&\geq \frac{1}{2}\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}(\mathbf{B}+\Gamma)\bar{\mathbf{d}} \\
&= \frac{1}{2}\mathbf{e}^T\bar{\mathbf{d}} \\
&= \mathbf{e}^T\mathbf{d}^{CP}
\end{aligned}$$

Since the oligopoly problem is convex (each subsidiary maximizes a concave objective over a convex set) $\mathbf{d}^{OP}$ is an oligopoly solution if and only if it satisfies the following quasi-variational inequality (see Bensoussan [11]):

$$(-\bar{p}_i + (\mathbf{B}+\Gamma)_i\ \mathbf{d}^{OP})\,(d_i - d_i^{OP}) \geq 0 \quad \forall d_i \in K(\mathbf{d}_{-i}^{OP}) \quad \forall i = 1,\ldots,n \tag{3.21}$$

where $K(\mathbf{d}_{-i}^{OP})$ is the feasible strategy space for subsidiary $i$ given the strategy $\mathbf{d}_{-i}^{OP}$ of the other subsidiaries. Since the constraint is active for every oligopoly equilibrium, the feasible strategy space is simply $K(\mathbf{d}_{-i}^{OP}) = \left\{d_i |\ d_i + \sum_{j\neq i}^{n} d_j^{OP} \leq C,\ d_i \geq 0\right\} = \{d_i|\ 0 \leq d_i \leq d_i^{OP}\}$.

In view of the feasible set $K(\mathbf{d}_{-i}^{OP})$, $\mathbf{d}^{OP}$ is an oligopoly solution if and only if for all $i$ such that $d_i^{OP} > 0$:

$$-\bar{p}_i + (\mathbf{B} + \Gamma)_i \mathbf{d}^{OP} \leq 0$$

On the other hand, $\mathbf{d}^{CP}$ satisfies the following KKT conditions for problem (3.7):

$$\begin{cases} -\bar{\mathbf{p}} + 2\mathbf{B}\ \mathbf{d}^{CP} + \mu\ \mathbf{e} - \boldsymbol{\lambda} = 0 \\ \mu\ (C - \mathbf{e}^T\mathbf{d}^{CP}) = 0 \\ \mathbf{e}^T\mathbf{d}^{CP} \leq C \\ \mu \geq 0 \\ \lambda_i\ d_i^{CP} = 0,\ \lambda_i \geq 0,\ \forall i = 1\ldots, n \end{cases}$$

For every $i$ such that $d_i^{CP} > 0$, $\quad -\bar{p}_i + 2B_i\mathbf{d}^{CP} + \mu = \lambda_i = 0.$

$$\begin{aligned} \Rightarrow -\bar{p}_i + (\mathbf{B} + \Gamma)_i \mathbf{d}^{CP} &\leq -\bar{p}_i + 2B_i\mathbf{d}^{CP} \\ &= -\mu_i \\ &< 0 \end{aligned}$$

$$\Rightarrow \mathbf{d}^{CP} \in S$$

□

When the energy consumption constraint is active for the centrally coordinated solution, with a carefully chosen allocation scheme $\mathbf{w}$, the reward-based incentive mechanism is able to induce the optimal subsidiary strategy for the company. In order to determine the weight allocation $\mathbf{w}$, the company can compute the optimal solution $\mathbf{d}^{CP}$ using problem (3.7), solve the KKT conditions of the oligopoly problem (3.3) at $\mathbf{d}^{CP}$ and use the normalized Lagrange multipliers $\boldsymbol{\lambda}$ of problem (3.3) for the weights. The company can thus coordinate the subsidiaries to achieve optimal profit using a decentralized method.

The question that remains is to understand the performance of this policy in terms of social welfare. How much does this policy hurt consumers? Will the regulator have an incentive to sue the company for monopolistic behavior? The answer is no. Compared to other oligopoly solutions this coordinated solution guarantees a decent social welfare.

**Theorem 3.14.** *When the joint capacity constraint* **is active** *for the centrally coordinated solution, the loss of social welfare is bounded by:*

$$TS(CP) \geq \frac{2}{3} TS(SMAX)$$

*Proof.* Refer to Appendix B.4. □

The loss of welfare in this case is at most 1/3 compared to the maximum loss of 1/4 in the unconstrained case (see Theorem 3.4).

As we just discussed (two paragraphs above), computing the appropriate weight allocation vector $\mathbf{w}$ requires the company to go through a complex optimization procedure. Moreover, the parent company needs to know the exact production costs and market characteristics of the subsidiaries to obtain the optimal weights. The subsidiaries have an incentive to distort the truth and not reveal their true parameters to obtain a bigger share of the profit. Collecting the accurate information might turn out difficult for the parent company. The subsidiaries may also complain that being subsidized at different rates is unfair. A simple and intuitive alternative for the parent company is to subsidize all the subsidiaries at the same rate. We next look at how this operationally simple procedure performs in terms of profit and welfare.

**Theorem 3.15.** *When the joint capacity constraint is active for the oligopoly problem, the loss of profit and social welfare under the uniform Nash equilibrium are bounded* *by:*

$$\begin{array}{rcl} \Pi(UNE) & \geq & \max\left\{\frac{2}{2+r\cdot(n-1)}; \frac{3}{4+r\cdot(n-1)}\right\} \Pi(CP) \\ TS(UNE) & \geq & \frac{5}{6}TS(CP) \\ TS(UNE) & \geq & \frac{5}{9}TS(SMAX) \end{array}$$

*where $r$ is the diversion ratio, $n$ is the number of subsidiaries.*

*Proof.* These results follow similarly to Kluberg and Perakis [76] since the Variational Inequality (VI) formulation for the uniform Nash equilibrium is the same as the VI for the disjoint constraints. Refer to Appendix B.4. □

Hence the uniform reward scheme leads to a limited loss of profit and welfare. With two subsidiaries, the loss of profit cannot exceed 1/3 for example. The loss of social welfare, on the other hand, is guaranteed to be below 44% independently of the number of subsidiaries or the market characteristics. These results show that the simple uniform reward scheme can extract a lot of the company profit in a decentralized manner without hurting consumers too much in the process. Note that these bounds originally established for oligopoly equilibria under general feasible convex sets might not be tight in our case. The uniform reward scheme performs potentially even better than the bounds.

## 3.7 Conclusions

This chapter studies the problem facing a company with multiple subsidiaries each of which needs to comply to a company-wide energy consumption target. While without a joint energy constraint, free (Cournot) competition only causes a limited (bounded) loss of company profit and loss of welfare, this chapter shows that in the presence of a joint constraint, the losses of profit and welfare due to competition can be arbitrarily bad even when the company only has two subsidiaries. The chapter then describes a reward scheme to coordinate the company's subsidiaries and reduce the losses of profit and welfare. It also points out particular company structures where the losses due to free Cournot competition are bounded. The company has several options in order to enforce compliance with the joint energy constraint: it can let the subsidiaries compete freely subject to the satisfaction of the constraint, it can offer a uniform subsidy to the subsidiaries per unit of energy saving to ensure compliance or it can offer a subsidiary specific energy subsidy with the goal to coordinate the production allocation. The chapter gives a potentially practical way of devising both the uniform and the subsidiary specific reward scheme. It shows the existence of a reward scheme that maximizes companywide profit and compares the performance of the different approaches in terms of companywide profit and in terms of social welfare: if a given scheme is too detrimental for consumers, a government regulator might intervene to prevent its implementation. Without any assumptions on the market characteristics, we show that while free competition can lead to extremely detrimental outcomes for the company as well as for society in general, the uniform reward scheme can ensure limited losses of profit and welfare. We then specialize the results to particular market structures. When the subsidiaries are fully symmetric, the uniform reward scheme is optimal from the point of view of the company as well as from the point of view of society. In this case, the uniform reward scheme should definitely be implemented. When the subsidiaries only present the same price potentials, the uniform reward scheme guarantees a loss of profit below 33.3% whereas free competition can take away up to 50% of the profit. In terms of social welfare, the uniform reward scheme does not hurt society compared to free competition, so the regulator will not try to prevent the company from implementing such a scheme. A possible next step in this research that goes beyond the scope of this chapter is to study these different schemes for a company facing multiple joint constraints.

# Chapter 4

# Loss of efficiency in deregulated electricity markets: a supply function equilibrium approach

## 4.1 Introduction

### 4.1.1 Motivation

The United States electricity market is a fascinating area for applied research. To begin with, its mere size is impressive. In 2009, the total US net electricity generation was roughly 4 billion Megawatthours (MWh). At an average retail price of 9.82 cents/kWh, the retail sale of electricity to consumers generated revenues of \$353.28 billion (see [132] for more details). These huge numbers highlight the tremendous impact, improvements in the electricity market can have. Reducing the electricity bill by 1% (by reducing energy consumption, by limiting electrical transmission losses or by driving electricity prices down for example) translates in savings of over \$3.5 billion. Another specificity of the electricity market, in contrast with other financial markets, is the complexity of its supply chain. The supply of electricity from producers to consumers is a four step process. It starts with the generating companies producing electricity and selling it in the wholesale market. Electricity is then transmitted at a high voltage (a few 100 Volts) from the generating locations to consumption areas. After being transformed to a much lower voltage (order of 10 Volts), it is then distributed to primary electricity customers (such as factories) and electricity distributors (such as NSTAR). Electricity is finally dispatched to individual consumers who purchase their electricity from the distributors. The electricity supply chain thus involves many intermediaries which raises the issues of coordinating them and allocating profits fairly among them. Moreover, electricity is not a storable good, so the generation and consumption of electricity must be balanced at every instant in time. Compared to a traditional supply chain, the service level here must be a 100%, all consumption must be satisfied, and excess electricity generation cannot be carried over as inventory. The electricity market finally differs from other financial markets by the absence of substitute to electricity and the necessity for a physical network to transport it. The electrical grid must satisfy a number of technical constraints. For example, lines in the grid have thermal capacities limiting the amount of electricity that can flow through them. Due to these specificities, the electricity market requires oversight: it cannot be settled via market processes only. In the United States, this oversight is provided by independent system operators (ISOs) which are independent organizations responsible for managing regional wholesale electricity markets and dispatching generators to insure a safe, feasible and reliable flow of electricity through the electrical grid. As an example, the New England ISO is responsible for the transmission of electricity in New England. It controls a transmission network of 8000 miles of high voltage lines and manages a generation capacity of approximately 31,000 MW. This generation capacity is distributed among more than 300 generators but the 20 largest generating companies own 27,000 MW out of the 31,000 MW (87%). The New England electricity market is thus a perfect example of the oligopoly nature of most US regional electricity markets.

This chapter is an attempt to understand the effects of free competition in such oligopolies. Since the 1890s, the production and dispatch of electricity has been a state controlled service. Under this regulated system, electricity was generated by large public utilities charging their marginal costs and it was dispatched by a State organization that had full control over the utilities. This regulated system had a single mission: keeping the lights on while "keeping costs down". As the Department of Energy puts it: "energy efficiency was a marginal consideration at best because energy was too cheap to be monitored". Starting around 1980 however, the deregulation process began with the PURPA act of 1978 and the EPA act of 1992. These policies put in question the necessity for a publicly run electricity system. They encouraged the proliferation of private generating companies and gave them access to the transmission network to allow them to compete with traditional utilities. Free competition between private generating companies is the basis of a deregulated market. After the California electricity crisis of 2000 though, a number of regions in the United States pulled the breaks on the deregulation process. Today, less than a third of the states in the US run a fully deregulated market. Legislators go back and forth, regulating and deregulating the electricity market. The problem is that they do not have a clear idea of the impact of deregulation because there is no rigorous and exhaustive study of the benefits and drawbacks of having a competitive electricity market.

In many areas, the electricity market is undergoing changes. Regulators, recognizing the need to modernize the power grid, initiated a redesign effort called Smart Grid. A Smart Grid is a grid applying "technologies, tools and techniques available now to bring knowledge to power, knowledge capable of making the grid work far more *efficiently...*" (DOE handbook on Smart Grid). With Smart Grid's recent advances, including the creation of demand response programs empowering consumers to curtail their consumption when electricity is too expensive, now is potentially a good time to move forward with deregulation.

This chapter provides a rigorous framework for the analysis of the deregulation process in electricity markets. Drawing on supply function equilibrium concepts, this chapter builds an improved model of competition in electricity markets. Compared to prior works (see next Section for references), our model reflects more accurately the role of the system operator and allows the formulation of network constraints. It enables the theoretical study and the simulation of the impact of deregulation on any given electricity market. Moreover, this chapter carries out a systematic analysis of the costs and benefits of deregulation for consumers, for generators and for society. In accordance with the actual objective of the system operator but in contrast with most of the literature (see next Section for references), this chapter uses social welfare as the central metric to evaluate the impact of deregulation. Under this metric, we compare a deregulated market with its regulated counterpart to give guidelines as to whether deregulation is beneficial or not.

### 4.1.2 Literature review

There is a large amount of research devoted to electricity markets. Broadly speaking, the literature can be split into four main categories. A number of papers focus on the decision problems facing generation companies such as whether to turn on or off their units (unit commitment problem, [130], [106], [126]), how to optimize their medium term electricity production scheduling ([105], [54], [51]) or how to plan for capacity expansion ([90], [53], [18]). A second stream of the literature attempts to forecast electricity demand (load forecasting, [23], [58]) or electricity prices ([9], [41], [55]) through statistical models. A third theme in the electricity literature is to model the security-constrained optimal dispatch problem solved by the system operator. While a number of papers represent the objective of the system operator as minimizing the cost of dispatch ([95], [88]), some papers recognize demand valuation for electricity and describe the goal of the system operator as to maximize social welfare ([4], [139], [22]). Maximizing social welfare is a fairer objective, as it values consumers' and generators' benefits equally, instead of focusing only on minimizing the price paid by consumers. In this chapter, we model the system operator as a maximizer of social welfare. In contrast to this chapter though, the literature modeling the system operator dispatch problem does not represent competition between the generators: it assumes the production costs of the generators known to the system operator and investigates the optimal dispatch given these costs. The last area of concentration in the literature is precisely devoted to understanding and designing competition in electricity markets. In this area, some papers model generators as competing through quantities (Cournot competition, [19], [21], [62], [63], [71]). The quantity competition model has the advantage that its equilibrium conditions are easier to establish and analyze and that it is computationally more tractable. However, this is an approximation, as in reality, demand for electricity is uncertain and generators must bid before knowing the demand. Yet, under the Cournot framework, generators' optimal quantity bids depend on the demand realization. To solve this dilemma, in practice each generator bids a menu of quantities and associated prices. To better reflect this reality, this chapter represents generators competition through a supply function equilibrium (SFE) model where each generator bids a function relating the amount of electricity it is willing to supply to the electricity price.

The supply function equilibrium model was first described in 1989 in the seminal paper by Klemperer and Meyer [75]. Since then, to better model the bidding process of the generators, a few papers have applied SFE concepts to electricity markets ([56], [2], [3], [6], [116], [32]). However, due to the complexity of the SFE model, some people claim these papers oversimplify the role of the system operator. They represent it as simply finding a clearing price that equates aggregate supply with demand. Unfortunately, this formulation is not flexible enough to include electrical network constraints. The true market clearing mechanism results from the optimization run by the system operator who looks for a dispatch that maximizes social welfare while respecting network constraints. The electricity market modeling literature is thus divided between Cournot models which imperfectly describe generators competition and SFE models which typically omit network constraints. This chapter presents a new supply function equilibrium formulation that explicitly represents the objective of the system operator and accounts for network constraints. The closest model to ours can be found in Hu & Ralph [69], but the emphasis there is on establishing conditions for existence and uniqueness of an equilibrium whereas the focus of this chapter is on evaluating the performance of deregulated markets.

In the presence of electrical network constraints, the best response problem of each generator to the supply function bids of its competitors is an MPEC (Mathematical program with equilibrium constraints). MPECs are in general hard to solve, see [83]. Finding a Nash equilibrium of supply functions for all the generators is thus an EPEC (Equilibrium problem with equilibrium constraints). Some papers modeling electricity markets through EPECS include [16] and [69].

After presenting the model, this chapter is devoted to evaluating the efficiency of deregulated markets. While the concept of efficiency is at the heart of the Smart Grid redesign plan, no precise definition of it exists in the electricity literature. Instead, to determine the efficiency of a market, two proxies are often analyzed. The first is the potential for *market power*, the ability of firms to manipulate equilibrium prices away from the optimum, which is usually detected through concentration measures (see FERC's reports [50], as well as [21], [66]). The second approach to measuring market inefficiency, is to simulate the cost of dispatch resulting from competition compared to the optimal, fully-regulated, cost of dispatch (see [56], [32]). While they provide interesting insights on the quality of a market structure, these metrics are incomplete representation of efficiency. Because they ignore the benefit consumers derive from receiving electricity, these metrics aim at minimizing the profit of generating firms. While the goal of regulators shouldn't be to maximize firms' profit, it shouldn't be to minimize it either. Both, consumers and producers are part of society and regulators should try to maximize their combined surplus. The issue is that electricity demand has traditionally been considered inelastic so it was hard to estimate consumers' valuation for electricity. Fortunately, this flaw of electricity markets is about to disappear. Indeed, one of the pillars of the Smart Grid project is a Demand Response Program which enables consumers to tailor their energy consumption based on real-time prices. Allowing consumers to actively participate in energy markets through demand response is essential to stabilizing prices and keeping the grid reliable. Demand response also provides us with a way of determining consumers' valuation of electricity (their willingness to pay). In contrast to the literature, this chapter uses social welfare (the sum of generators' profit and consumer surplus) as the central metric to evaluate the performance of deregulation. [136] analyzes the impact of SFE competition on social welfare but, unlike ours, this chapter models generators facing private random production costs and deterministic demand (instead of deterministic production costs and random demand) and does not include network constraints. While this chapter focuses on how the market characteristics (supply and demand elasticity, number of generators) affect social welfare, [136] is concerned with the impact of the production costs' randomness.

We compare SFE competition with the fully regulated setting. The comparison methodology adopted in this chapter draws on the concept of *price of anarchy* as described in the seminal paper by Koutsoupias and Papadimitriou [77]. The idea is to evaluate the ratio between the worst case equilibrium (here supply function equilibrium) resulting from competition and the coordinated (or regulated) solution. The idea of studying the loss of efficiency is common in the field of transportation ([31], [115]) and in the field of supply chain management ([29], [104]) for example. The price of anarchy is also measured in the context of pure oligopolistic competition such as Bertrand competition [48] or Cournot competition [60]. This chapter is the first application of the price of anarchy methodology to the study of electricity markets using the SFE model.

### 4.1.3 Contributions and Outline

This chapter brings a number of contributions to the literature. First, it **generalizes the SFE model** in electricity markets to account explicitly for the role of the system operator and to include network constraints. While traditional SFE models of electricity markets are typically single level Nash games between the generators subject to the market rule of matching supply and demand, our model is a bi-level Stackelberg game where each generator anticipates the dispatch of the system operator while taking the bids of its competitors as given (i.e. solves a best response problem) and where the system operator runs an optimization to determine the dispatch. Our model is not only a more accurate description of the bidding process, but it also allows the SFE model to incorporate electrical network constraints (in the optimization of the system operator) and it does not assume that the system operator knows the exact production costs of the generators. Second, this chapter **emphasizes the importance of social welfare** as a metric to include in the objective function of the system operator dispatch and also as an evaluation tool to measure the performance of deregulation in a given market. To model SFE competition, we come up with a proxy of social welfare $ProxyTS(.)$ that allows the system operator to optimize dispatch without knowing the production costs of the generators. With this model, this chapter **carries a systematic analysis of the costs and benefits of deregulation** for the generators (through profits), for the consumers (through consumer surplus) and for society as a whole (through social welfare). To quantify the performance of deregulation, this chapter **introduces two new market metrics**: the unit-less relative elasticity of demand versus supply (denoted by $z$) and the number $r$ of "significant" generators in a market where generators vary in size. Finally, this chapter simulates competition on two realistic electrical networks (see Section 4.5 for more details) and **compares the performance of deregulation on these systems to our theoretical bounds**.

The chapter shows that the number of generators, or more exactly the number of "significant" generators $r$, is a critical factor in determining whether deregulation is beneficial or not. With only one generator supplying electricity, deregulation is certainly not advisable. In a deregulated market, the monopolistic generator can exercise a lot of market power, drive its profit up while harming consumers and decreasing overall social welfare. However, this chapter shows that this effect disappears quickly as competing generators are introduced in the market. Even with only two generators, consumers can lose at most 32% of their surplus and the welfare of society cannot decrease by more than 3% compared to the regulated setting. With three generators or more competing, the loss of consumer surplus improves further and is bounded by 20% and that of social welfare by 1%. The numerical simulations carried out on our two network systems corroborate these results: deregulated SFE competition is able to extract 99.9% and 99.2% of the maximum social welfare in the 30-bus and 5-bus system respectively.

We begin in Section 4.2 by defining the notion of social welfare and by using it to introduce a new Stackelberg SFE model of electricity markets. We then show the performance of this SFE competition on a market with symmetric and asymmetric generators without network constraints in Section 4.3 and 4.4 respectively. We finally simulate the performance of our model on two electricity test-systems, a 30-bus and a 5-bus system, in Section 4.5. We compare our simulation results with the regulated policy and with Cournot competition.

## 4.2 The supply function equilibrium (SFE) model

This chapter models and analyzes competition between generators supplying electricity to consumers through an electrical grid. We assume that $n$ firms compete in the market: $i = 1, \ldots, n$. As is the case in practice, the number $n$ of generators is relatively small and the decisions of each of these generators can influence the final price of electricity: these generators are atomic players in the market. In contrast with other markets, the electricity market cannot be fully autonomous. The absence of substitute to electricity for consumers and the necessity to maintain a safe, feasible and reliable flow through the electrical grid require the presence of a regulator. This role is played throughout the United Stated by independent institutions called independent system operators (ISO) in charge of administering the market and monitoring the grid for a specific region. The wholesale market analyzed here serves large consumers composed of electricity distributors (such as NSTAR in New England) and factories. A diagram of the electricity market is shown in Figure 4-1.

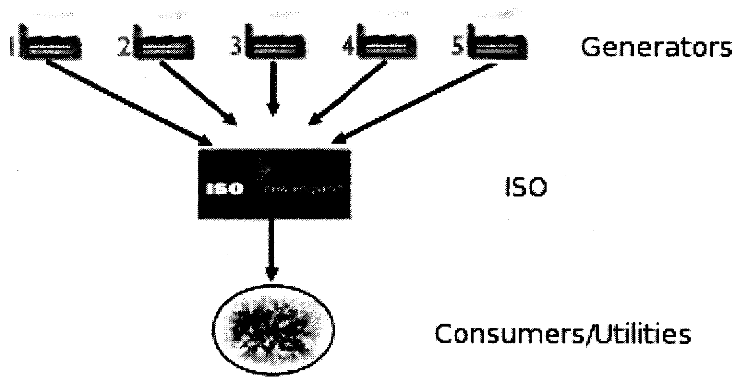


Figure 4-1: Electricity market diagram.

To run the market, the system operator traditionally used to forecast consumption for every hour of the day and to take this forecast as fixed demand. It would then accept bids from the generators, and would dispatch them to create a feasible electricity flow satisfying demand at minimum cost. This process is evolving though as a lot of emphasis is placed on demand response programs which allow consumers to curtail their demand when electricity is expensive. Contrary to [2], [88] and [116] for example, consumption is no longer viewed as a fixed quantity but as a function of electricity price. Accordingly, we represent consumption in our market through an aggregate demand curve $D(p)$: demand is elastic. $D(.)$ is taken to be a strictly decreasing and concave function on its support (i.e. some interval $(0, \hat{p})$, with $\hat{p}$ such that $D(\hat{p}) = 0$). In fact, while the generating firms bid once a day in the electricity market, electricity demand fluctuates throughout the day. We represent this demand variation by the uncertainty parameter $\epsilon$; demand is described through the function $D(p, \epsilon)$. We do not make any assumption on the distribution of uncertainty parameter $\epsilon$ except that it has a non-negative support. Our supply function equilibrium model is well defined under this general demand form (see [75] for proof) but to simplify calculations in this chapter, we consider a linear demand model: $D(p, \epsilon) = -m\, p + \epsilon$. $\epsilon$ is the intercept, i.e. the maximum potential demand (at price 0), and $m$ is the demand sensitivity to the electricity price $p$. Assuming a linear demand is common in the literature, see for example [131], [56], [21], [6] and [30].

Generator $i$ incurs a global production cost $C_i(q_i)$ to produce quantity $q_i$. Cost $C_i(.)$ is assumed increasing and convex as generator $i$ will first turn on its most efficient (cheapest per unit) plants to produce electricity, resorting to using more and more expensive plants as it reaches its total production capacity. These assumptions on the production costs are enough for the supply function equilibrium model developed below to be meaningful (see [75]) but in the following sections, we will restrict attention to quadratic production costs. As in [131], [56], and [57], each firm thus faces a production cost: $C_i(q_i) = b_i q_i + 1/2 c_i q_i^2,\ i = 1, \ldots, n$. We do not model the constant terms in the quadratic production costs (i.e. $a_i = 0$ in $C_i(q_i) = a_i + b_i q_i + 1/2 c_i q_i^2$) as such constants do not influence the pricing or production decisions of the generators (these are fixed, sunk costs independent of the generators' decisions).

Our supply function equilibrium model builds on the seminal paper [75] by Klemperer and Meyer. Each generator submits a supply function $S_i : [0, \hat{p}) \to (-\infty, \infty)$ to the market with the goal to maximize its profit from selling electricity. Following [75], most of the SFE literature on electricity (see Section 4.1.2 for references) defines the market mechanism as simply matching supply with demand. Under this market mechanism, after all the firms choose their supply functions, an equilibrium price $p^*$ is established such that $D(p^*, \epsilon) = \sum_{i=1}^{n} S_i(p^*)$ and generators supply $q_i^* = S_i(p^*)$ ($p^*$ is a scalar denoting the price of electricity which is a homogeneous good). If firm $i$ knew the supply functions submitted by its competitors, it would then aim at maximizing its profit (over the scalar $p$) given its residual demand $S_i(p) = D(p, \epsilon) - \sum_{j \neq i} S_j(p)$ (i.e. solve the best response problem):

$$\max_p \; p\Big[D(p, \epsilon) - \sum_{j \neq i} S_j(p)\Big] - C\Big(D(p, \epsilon) - \sum_{j \neq i} S_j(p)\Big)$$

This optimization leads to a point $(p_\epsilon^*, q_{i,\epsilon}^*)$ for each $\epsilon$, with $q_{i,\epsilon}^* = D(p_\epsilon^*, \epsilon) - \sum_{j \neq i} S_j(p_\epsilon^*)$. The supply function $S_i(p)$ is an optimal response to the supply function of its competitors if it passes through all these optimal points, namely $q_{i,\epsilon}^* = S_i(p_\epsilon^*)$, for all $\epsilon$.

However, this model has two major drawbacks. For one, it forgets about the role played by the independent system operator (ISO) whose actual goal is not only to match supply with demand but also to maximize efficiency (in terms of social welfare). Moreover, under the simple market mechanism that matches supply with demand, it is not possible to integrate the electrical network constraints into the model. This limits the model to a high level economic abstraction of actual electricity markets. Network effects play a major role in the electricity price and dispatch decisions of the system operator. In order to address these drawbacks, we, as a result, define a new market mechanism which, combined with the supply function equilibrium framework, provides a complete representation of electricity markets.

In an abstract manner, the market can be represented as a process which takes the supply functions $S_i(.)$ of all the generators as input and outputs the equilibrium price $p^*$ and the generation quantities $q_i^*$ for all the generators. We denote this process by: $(p^*, \mathbf{q}^*) = \mathbf{Market}(S_1(.), \ldots, S_n(.))$ where $\mathbf{q}^*$ is the vector of generation quantities $q_i^*$'s. One possible implementation of this market mechanism is simply to match supply and demand as described above. We will show below a more realistic implementation of $\mathbf{Market}(.)$. Given the market price $p^*$ and the dispatch quantity $q_i^*$, generator $i$ receives a profit of: $p^*\ q_i^* - C_i(q_i^*)$. In a general sense, a supply function equilibrium is characterized by each generator $i$ solving a best response problem and deciding its supply bid $S_i(.)$ through the optimization:

$$\begin{aligned} \max_{S_i(.)} \quad & p^* \cdot q_i^* - C_i(q_i^*) \\ \text{s.t.} \quad & (p^*, \mathbf{q}^*) = \mathbf{Market}(S_i(.), \mathbf{S}_{-i}(.)) \end{aligned} \tag{4.1}$$

taking the supply functions of the competitors $\mathbf{S}_{-i}(.)$ as given.

In contrast to [75], our model represents the market mechanism $(p^*, \mathbf{q}^*) = \mathbf{Market}(S_1(.), \ldots, S_n(.))$ as the result of an optimization run by the ISO.

As explained above, the goal of the system operator is to maximize the welfare of society (denoted by $TS$ for total surplus) while respecting the bids of the generators and maintaining a feasible electricity flow on the grid. The system operator first observes the realization of demand uncertainty $\epsilon$. It then solves:

$$
\begin{array}{rclll}
(p_\epsilon^*, \mathbf{q}_\epsilon^*) & = & \underset{p,\mathbf{q}}{\text{argmax}} & TS_\epsilon(p, \mathbf{q}) & \\
& & \text{s.t.} & \left\{ \begin{array}{l} \sum_{i=1}^n q_i = D(p, \epsilon) \\ 0 \le q_i \le S_i(p) \\ \Phi q \le \Delta \end{array} \right. & \begin{array}{l} \text{(Supply = Demand)} \\ \text{(Supply bids)} \\ \text{(Net. Const. - DC Approx)} \end{array}
\end{array}
\tag{4.2}
$$

We use the traditional DC approximation of electricity flows which allows us to model grid constraints as linear constraints. See [74] for references.

Under our concave (linear) demand assumption, our convex (quadratic) production costs and some mild continuity and monotonicity assumptions on the set of supply functions, problem (4.2) has a unique solution for every realization of the uncertainty $\epsilon$ (see [69] for a proof). Let us denote problem (4.2) by the optimal dispatch function: $(\mathbf{q}_\epsilon^*, p_\epsilon^*) = \text{OPTIDISP}_\epsilon(S_1(.), \ldots, S_n(.))$.

In this model, the sequence of events is as follows. The generators $i = 1, \ldots, n$ submit their supply functions $S_i(.)$ first, then demand is realized and finally the ISO dispatches the generators according to $\text{OPTIDISP}_\epsilon(.)$.

Aware that the ISO dispatches according to (4.2) and taking the bids $\mathbf{S}_{-i}(.)$ of the competitors as given, each generator aims at maximizing its profit across all realizations of the demand. For ease of exposition, we assume in this chapter that all demand scenarios are uniformly likely to occur. Each generator $i$ hence chooses its bid to maximize:

$$
\begin{array}{rl}
\underset{S_i(.)}{\max} & \sum_\epsilon p_\epsilon^* \, q_{i,\epsilon}^* - C_i(q_{i,\epsilon}^*) \\
\text{s.t.} & (p_\epsilon^*, \mathbf{q}_\epsilon^*) = \text{OPTIDISP}_\epsilon(S_i(.), \mathbf{S}_{-i}(.)) \quad \forall \epsilon
\end{array}
\tag{4.3}
$$

The model easily extends to non-uniform probability distributions of the stochastic demand. In that case, the profit under each demand realization must be multiplied by the probability of realization $\mu(\epsilon)$ of that demand scenario in the objective function of (4.3). To keep exposition simple, this chapter presents only the case of uniform demand distribution but the insights are similar under any distribution.

To determine the optimization of the system operator (4.2), we need to define the measure of social welfare: $TS(p, \mathbf{q})$. The idea of social welfare is to aggregate the benefit of the generating companies with the benefit of consumers to produce a solution that is fair and rewards all market participants. We have already described the benefit of the generating companies in this market. Together the generating companies make a profit:

$$\text{Profit}(p, \mathbf{q}) = \sum_{i=1}^{n} [p\ q_i - C_i(q_i)] \tag{4.4}$$

To evaluate the benefit of consumers, we precisely use the newly adopted framework of demand response. As described above, demand is modeled through a demand function $D(p, \epsilon) = -m\ p + \epsilon$. We draw upon the classical economic notion of consumer utility (see [135] Chap, 6 for reference), which states that the demand function $D(p, \epsilon)$ is in fact derived from consumers having an aggregate utility function: $U_\epsilon(d) = \frac{\epsilon\, d}{m} - \frac{d^2}{2m}$. This utility function represents the benefit (evaluated in \$) for the consumers of using $d$ MW of electricity. Consumer surplus is then defined as the difference between the utility derived by consumers from their electricity use and the price paid by consumers to purchase this electricity:

$$CS_\epsilon(p, d) = U_\epsilon(d) - p\ d \tag{4.5}$$

Social welfare is then simply defined as the sum of the generators' profit and the consumer surplus:

$$\begin{aligned} TS_\epsilon(p, \mathbf{q}) &= CS_\epsilon\Big(p, \textstyle\sum_i q_i\Big) + \textstyle\sum_i p\ q_i - C_i(q_i) \\ &= U_\epsilon\Big(\textstyle\sum_i q_i\Big) - \textstyle\sum_i C_i(q_i) \end{aligned} \tag{4.6}$$

At this stage, our model still has one major problem. The system operator is supposed to solve problem (4.2), where the objective function $TS$ is defined in (4.6). In particular to calculate $TS$, the system operator needs to know the production costs $C_i(.)$, $i = 1, \dots, n$ of all the generators. However, in a deregulated market, one of the key issues is that the system operator precisely does not know these costs. Indeed, the very basis of a deregulated market is that an efficient electricity price is achieved by letting generators compete to supply electricity and not by controlling them. While under the regulated system, the State monitors the actual production costs of its utilities, in the deregulated market the generators are private companies and their production costs are private, unknown to the system operator. Under a deregulated market, the only information the system operator has on generators is their bids $S_i(.)$. It must rely on these bids to make dispatch decisions.

To resolve this difficulty, we assume that the system operator precisely uses generators' bids $S_i(p)$ to infer a proxy of their production costs. Given a fixed electricity price $p$, a generator with production cost $C(q)$ would decide its supply quantity $q$ by solving:

$$q = \underset{\tilde{q} \geq 0}{\operatorname{argmax}} \quad p\, \tilde{q} - C(\tilde{q})$$

Under convex production costs, the optimality conditions lead the generator to choose a production quantity $q$ satisfying: $p = C'(q)$. The supply function submitted by the generator, on the other hand, requires him to produce a quantity $q = S(p)$ at price $p$. Inverting this relation gives $p = S^{-1}(q) = C'(q)$. The function $\widetilde{C}(q) = \int_{u=0}^{q} S^{-1}(u)du$ is therefore a meaningful proxy of the production cost of the generator. In fact, $\widetilde{C}(q)$ exactly matches the production cost of the generator, when he is fully rational, has unlimited capacity and considers himself a price-taker. We denote the social welfare that uses this proxy for costs as described above by:

$$ProxyTS(p, \mathbf{q}) = U_\epsilon\Big(\sum_i q_i\Big) - \sum_i \widetilde{C}_i(q_i)$$

To summarize, our supply function equilibrium model is defined by the generators simultaneously solving problem (4.3), in which OPTIDISP$_\epsilon$ is the outcome of optimization problem (4.2) where the objective of the system operator is $ProxyTS(p, \mathbf{q})$ instead of $TS(p, \mathbf{q})$. This is actually close to what is happening in practice. The generators are bidding supply functions and the system operator chooses the electricity dispatch by solving problem (4.3). Yet, while our model reflects current practice, it is not commonly adopted in the literature due to its complexity (see Section 4.1.2 for references). Existence and uniqueness of a solution to such a bi-level supply function model are shown in [69].

To keep computations tractable and the presentation of the chapter clear, we adopt a linear supply function equilibrium model. That is, we consider generators bidding linear supply functions of the form: $S_i(p) = \beta_i(p - \alpha_i), \quad i = 1, \ldots, n$. In problem (4.1), each generator $i$ needs to determine the coefficients $\alpha_i$ and $\beta_i$ of its bid $S_i(.)$. Restricting attention to the space of linear functions is common in the SFE literature ([75], [131]), particularly in modeling electricity markets ([56], [57]). Baldick et al. [6] provide a detailed discussion on the choice of linear supply functions for electricity markets. [7] justifies to concentrate attention on linear supply function equilibria because other equilibria are unstable.

Generalizing the supply function equilibrium model from its original formulation by Klemperer and Meyer [75] to a model that reflects the actual objective of the system operator and accounts for electrical network constraints is the first contribution of this chapter. We reformulated the dispatch of the ISO as an optimization problem that accounts for network constraints. We defined a proxy of social welfare that is used by the system operator to choose the optimal dispatch. This is accomplished above. The remainder of this chapter is devoted to evaluating the performance of the SFE model in comparison with the centrally regulated setting and the quantity (Cournot) competition model.

Under the traditional regulated utility model, the system operator monitors the production costs of the participating generators and dispatches them in order to maximize the welfare of society. In contrast with the SFE model, under the regulated setting, the generators do not bid, but rather, the ISO can directly use $TS(p, \mathbf{q})$ in its optimization because it knows the generators' production costs. The optimization of the ISO can therefore be described as:

$$
\begin{array}{rcl}
(\widetilde{p}_\epsilon, \widetilde{\mathbf{q}}_\epsilon) & = & \underset{p,\mathbf{q}}{\text{argmax}} \quad TS_\epsilon(p, \mathbf{q}) \\
 & & \text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} q_i = D(p, \epsilon) \\ 0 \leq q_i \\ \Phi q \leq \Delta \end{cases}
\end{array}
\tag{4.7}
$$

We study the performance of the supply function equilibrium model from the perspective of consumers (as measured by consumer surplus (4.5)), from the perspective of generators (as measured by aggregate profit (4.4)) and from the perspective of the ISO (as measured by social welfare (4.6)). The value of the SFE model for the different market participants (Profit$(Dereg.)$, $CS(Dereg.)$ and $TS(Dereg.)$) varies widely depending on the market structure (such as the number of participating generators) and the configuration of the electrical grid. So instead of looking at these absolute valuation measures, the SFE performances are benchmarked against the performances of the centrally regulated system. We therefore compute the ratios $\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)}$, $\frac{CS(Dereg.)}{CS(Centr.)}$ and $\frac{TS(Dereg.)}{TS(Centr.)}$. The goal of the following sections is to evaluate these ratios on various market structures and electrical networks, in order to show how consumers, generators and society are affected by the deregulation process (the transition from a centralized system to SFE competition). Sections 4.3 and 4.4 investigate markets with respectively symmetric and asymmetric generating firms and no electrical constraints. Section 4.5 simulates SFE competition on two realistic electrical grids with network constraints.

## 4.3 Unconstrained symmetric generators

We begin the analysis by considering a market without electrical network constraints. A good discussion on the meaning and the formulation of electrical network constraints can be found [74]. Due to the already complex analysis of the SFE model by itself, most of the economics literature, including the electricity market literature ([2], [6], [56], [57], [131]), has avoided to include constraints in its models. Such models represent networks where the thermal (flow) capacity of the electrical lines are not a major bottleneck and do not influence the dispatch decisions of the system operator. This is obviously a limiting assumption but it is a good first cut assumption that helps the understanding of the SFE model. While the SFE model without network constraints has been the subject of previous papers (cited above), our analysis is new for two reasons. It is the first paper to carry a systematic comparison between SFE competition and the centrally regulated system from the perspective of consumers and generators. It is also the first paper to focus on social welfare ($TS$ as defined in (4.6)) to evaluate the performance of deregulation. In contrast, the electricity market literature, has been investigating electricity prices, generators' profit and the exercise of market power in order to assess deregulated models. Nevertheless, one may argue that these measures are not fair to the generators as we discussed in the introduction.

Even without network constraints, equilibria in supply functions are not trivial to compute as is shown in Section 4.4. We analyze in this section an oligopoly market of $n$ symmetric generating firms. The firms all face the same quadratic production costs: $C(q) = bq + 1/2\ c\, q^2$. $b$ is the constant unit cost, $c$ represents the increase in marginal production cost due to machine capacities: that is, the most cost-effective machines are used first, then the more expensive ones. The firms being symmetric, or identical, they compete with each other on an equal basis. As derived in Appendix C.1, SFE competition is partially truth revealing as it forces generators to bid a reservation price $\alpha_i$ (the supply bids are $S_i(p) = \beta_i(p - \alpha_i)$) equal to the linear coefficient $b$ of their production costs. This property is known as incentive compatibility in the economics literature, see [6]. Under our symmetry assumption, all the generators will bid $\alpha_i = b$ and it can easily be shown that there is no loss of generality in setting $b = 0$. Section 4.3 follows this convention.

We investigate next the effects of deregulation on $CS$, Profit and $TS$, for a monopoly market, a duopoly market and a general oligopoly market. Under our assumptions, all the relevant ratios ($\frac{\text{Profit}(\textit{Dereg.})}{\text{Profit}(\textit{Centr.})}$, $\frac{CS(\textit{Dereg.})}{CS(\textit{Centr.})}$ and $\frac{TS(\textit{Dereg.})}{TS(\textit{Centr.})}$) are independent of the realization of the demand uncertainty $\epsilon$. The uncertainty level $\epsilon$ affects the regulated and the deregulated settings in a proportional way, and as a result, it does not affect the ratios . Moreover, remarkably enough, these ratios depend on the parameter $m$ of the demand function and the parameter $c$ of the production costs only through a single unit-less parameter $z$ (see Appendix C.1 for proof).

We define the parameter $z = m\ c$ as the relative elasticity of demand versus supply. $z \to 0$ when $m \to 0$ which means demand is inelastic compared to supply ($D(p) = \epsilon$). Conversely, $z$ tends to infinity with $m$ when demand is extremely elastic. Since $C'(q) = c\ q$ and as established in the previous section, generators want to produce at marginal cost $S(p) \approx (C')^{-1}(p) = \frac{1}{c}\ p$, the coefficient $z$ can be interpreted as the ratio of demand elasticity $m$ and supply elasticity $\frac{1}{c}$. It makes sense that $z$ is the meaningful measure of elasticity, as demand is only elastic in comparison with supply.

Before getting into the analysis of the performance ratios and the effect of the number of competing generators on these ratios, we examine the impact of deregulation on electricity price and quantity. No matter the number of generators, electricity price is higher and production quantity is lower under the deregulated setting than under central regulation.

**Proposition 4.1.** *When symmetric generators compete through a supply function bidding mechanism, they always achieve a greater combined profit than under the centrally regulated solution. They charge a higher electricity price and produce less* *electricity.*

$$p^*(Dereg) \geq p^*(Centr.), \quad q(Dereg.) \leq q(Centr.), \quad \text{Profit}(Dereg.) \geq \text{Profit}(Centr.)$$

*Proof.* See Appendix C.1. □

The idea is pretty intuitive. In the absence of regulation, generators compete freely and they can exercise some market power by submitting supply bids that understate the amount of electricity they can produce profitably at a given price. Doing so, the generators will be dispatched to produce less electricity, they will drive electricity prices up and derive higher profits.

### 4.3.1 The case of a monopolistic generator

Even with a single generator supplying electricity, deregulation has a significant impact on the market participants. In fact, as this section shows, deregulation has the biggest impact in the case of a monopolistic generator. Whereas under the regulated utility model the system operator knows the generator's costs and optimizes the dispatch to benefit society as a whole, under free competition the generator can exercise a lot of market power by bidding to supply low levels of electricity as a function of price. Since there are no other generators to supply electricity in its place, the monopolistic generator can limit supply, push the electricity price up and drive its profit higher.

While the increase in price under deregulation might be expected because of the lack of vertical coordination, it is not a consequence of the well documented double-marginalization process that occurs in a decentralized vertical supply chain (see [121]). In contrast to traditional vertical supply chains, in this chapter, the intermediary, the system operator, does not have as objective to maximize its own profit. Rather, it seeks to maximize social welfare (including the generators profit). This chapter shows that even when the intermediary's goal is to coordinate the whole system, the lack of coordination due to generators competition still pushes the electricity price up.

In Figure 4-2, we plot the various performance ratios (consumer surplus, profit and social welfare) that are derived in closed form in Appendix C.1 as functions of the unit-less relative elasticity $z$. The top curve in Figure 4-2 represents the profit ratio $\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)}$. Notice that the curve stays above 1, signaling that the generator is always better off in the deregulated setting than under central regulation. Even though in the deregulated market there is a vertical game between the generator and the system operator, the generator is better off because there is no double-marginalization effect here. While the generator and the system operator have divergent objectives, these do not directly compete with each other. When demand is very inelastic ($z = 0$), a monopolistic generator can exercise a lot of market power and it can make infinitely more profit under the deregulated setting than under central regulation. For example, when demand and production have the same price elasticity ($z = 1$), a monopolistic generator makes 40% more profit in the absence of regulation.

We now look at the performance of the SFE model in terms of social welfare.

**Theorem 4.1.** *When a single generator is free to bid its supply function to the system operator, society can lose at most 25% (100-75%) of its welfare compared to the situation where the market is centrally regulated.*

$$\frac{TS(Dereg.)}{TS(Centr.)} \geq 75\%$$

*Proof.* See Appendix C.1. □

Under the regulated setting, the system operator controls all the generators and dispatches quantities precisely with the goal to maximize social welfare. It is thus clear that the regulated setting is optimal in terms of social welfare and that the social welfare ratio $\frac{TS(Dereg.)}{TS(Centr.)}$ is less than 1. The goal in analyzing the social welfare ratio is to measure how large the loss of social welfare caused by deregulation is. If this loss is not too large, it might be offset by the savings made in allowing a free competitive market as opposed to paying the system operator to monitor the production costs of the generating companies. Even with a single generating firm, Theorem 4.1 establishes that the loss of social welfare due to deregulation cannot be greater than 25%. The middle curve in Figure 4-2 represents the social welfare ratio as a function of the unit-less parameter $z$. When demand and production have the same price elasticity ($z = 1$), the loss of social welfare is only 10%.

We finally adopt the perspective of consumers.

**Theorem 4.2.** *When a single generator is free to bid its supply function to the system operator, consumers lose up to 75% (100-25%) of their surplus compared to the situation where the market is centrally regulated.*

$$\frac{CS(Dereg.)}{CS(Centr.)} \geq 25\%$$

*Proof.* See Appendix C.1. $\square$

The bottom curve in Figure 4-2 represents the consumer surplus ratio as a function of the unit-less parameter $z$. While bounded (it can be no more than 75%), the loss of consumer surplus here is large independently of $z$. The reason is that with only a single generator and price-taking consumers, the generator is able to draw a lot of the social wealth towards its own profit, leaving very little benefit for the consumers. A fairness argument would suggest not to deregulate in such a market. Although the loss of social welfare is reasonable, there is a huge, unjustified, transfer of wealth from the consumers to the generator. This is precisely the behavior, market power studies are aimed to detect and prevent.

### 4.3.2 The case of two generators

As in [131] and [56], we consider the duopoly case separately. The reason we examine it separately here is to point out the remarkable performance improvement for consumers in going from a monopoly to a duopoly market. Here too, deregulation allows the two generators to exercise some market power by submitting supply functions

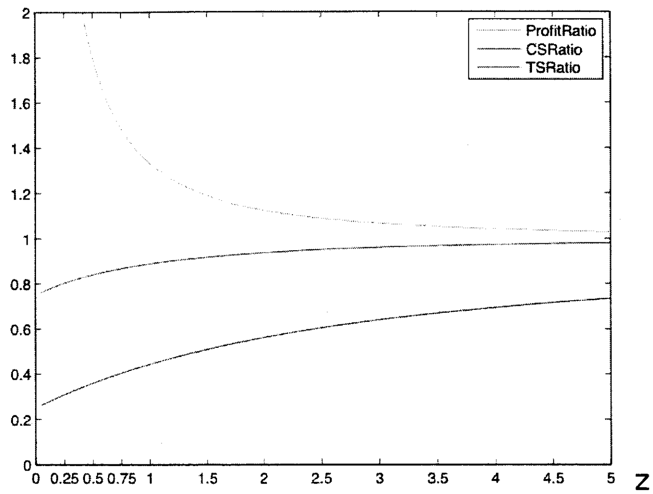


Figure 4-2: Welfare, surplus and profit ratios for a monopolistic generator as functions of the unit-less cost parameter $z$.

that understate the amount of electricity they could produce profitably at a given price. While the two generators are competing with each other, they are still able to limit supply, push the electricity price up and increase their profit.

In Figure 4-3, we plot the performance ratios (surplus, profit and welfare) derived in Appendix C.1 in closed form, as functions of the relative elasticity $z$. The top curve in Figure 4-3 represents the ratio of global profit $\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)}$. The curve stays above 1, signaling that the generators are always better off in the deregulated setting than under central regulation. For example, when demand and production have the same price elasticity ($z = 1$), a generator in a duopolistic market makes 60% more profit in the absence of regulation.

**Theorem 4.3.** *With two symmetric generators competing through a supply function bidding mechanism, society can lose at most 3% (100-97%) of its welfare and consumers can lose at most 32% (100-68%) of their surplus compared to the centrally regulated solution.*

$$\frac{TS(Dereg.)}{TS(Centr.)} \geq 97\% \qquad \frac{CS(Dereg.)}{CS(Centr.)} \geq 68\%$$

*Proof.* See Appendix C.1. □

The loss of welfare (similarly profit and consumer surplus) is defined as 1 minus the minimum welfare (or profit or surplus) ratio. The addition of a only one generator to a monopolistic market almost eliminates the entire loss of welfare, reducing it from 25% to 3%. The middle curve in Figure 4-3 represents the social welfare ratio as a function of the unit-less parameter $z$.

The bottom curve in Figure 4-3 represents the consumer surplus ratio as a function of the unit-less parameter $z$. There is a net improvement for consumers in going from a monopolistic to a duopolistic market. The two generators are still able to draw a substantial amount of social wealth towards their profit. Consumers can still be deprived of a lot of surplus, up to 32%, but that is a lot less than the 75% potential loss under a monopolistic market.

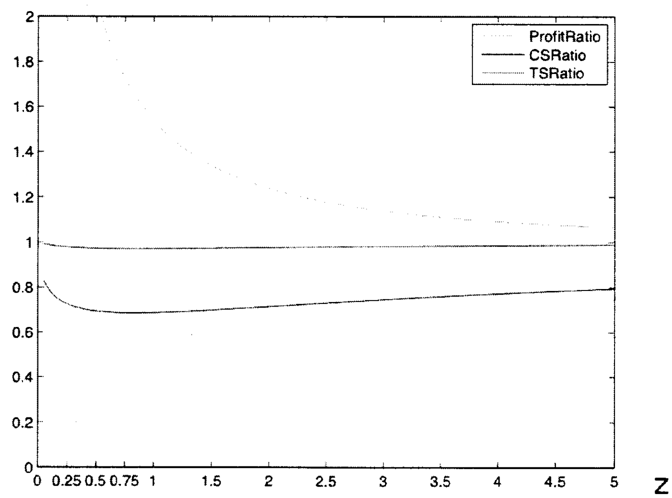


Figure 4-3: Welfare, surplus and profit ratios for two symmetric generators as functions of the unit-less cost parameter $z$.

### 4.3.3 The case of many generators

The insights derived in the monopoly and duopoly cases still hold with many generators. Competing generators are still better off under free competition than under central regulation. While the generators are competing with each other, they are still able to exercise market power and drive profits higher.

Figure 4-4 represents the ratio of generators' profit $\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)}$ for markets with various number of generators. The curves all stay above 1, as generators are always better off in the absence of regulation. The profit ratio is monotonically decreasing as a function of the demand elasticity $z$ ($z$ increases as demand becomes more price sensitive relative to supply). The intuition behind this monotonicity is that when demand is inelastic, the generators have more power relative to consumers to drive prices and profit up by limiting supply. However, these curves do not exhibit any monotonicity pattern as a function of the number of generators. When demand is inelastic ($z$ small), a small number of generators can extract more profit than can many. The contrary is true when demand is very elastic.

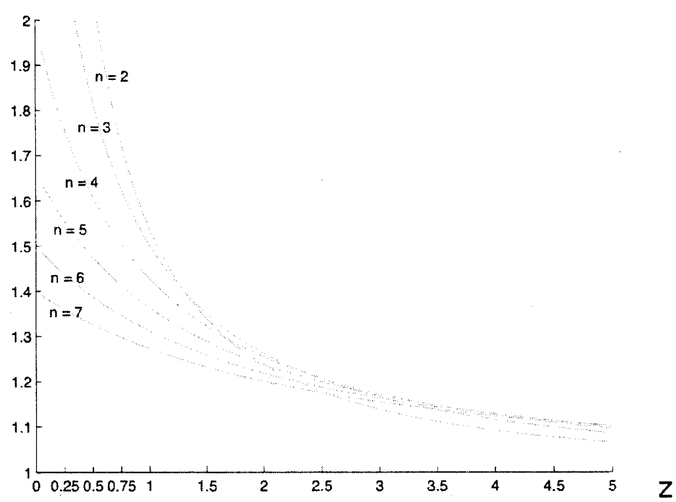


Figure 4-4: Profit ratios for $n$ symmetric generators ($n$ ranging from 2 to 7) as functions of the unit-less cost parameter $z$.

The next theorem examines the impact of deregulation on social welfare for symmetric oligopoly markets.

**Theorem 4.4.** *With $n$ symmetric generators ($n \geq 3$) competing through a supply function bidding mechanism, society can lose at most 1% of its welfare compared to* *the centrally regulated solution.*

$$\frac{TS(Dereg.)}{TS(Centr.)} \geq \Phi_1(n) \geq 99\% \qquad \text{for } n \geq 3$$

*Proof.* See Appendix C.1. □

Figure 4-5 depicts the social welfare ratio $\frac{TS(Dereg.)}{TS(Centr.)}$ as a function of the demand elasticity $z$ for oligopoly markets with various number of generators. As explained at the beginning of the section, our performance ratios only depend on the demand function and the production costs through the unit-less parameter $z = m\ c$. Let us denote the curves of Figure 4-5 by $\Phi_1(n, z)$. These functions are obtained in closed form in Appendix C.1 by calculating the social welfare under SFE competition and under the regulated model. Observe that the loss of social welfare is decreasing with the number of generators uniformly across all values of the market parameter $z$. This means that the adverse effect of deregulation on the welfare of society becomes small when many generators participate in the market. This coincides with the common intuition that deregulation is only beneficial when the market structure is competitive enough. Moreover, the functions $\Phi_1(n, z)$ are unimodal and admit a minimium at $z = -n + 2\ \sqrt{n^2 - n}$. The bound $\Phi_1(n)$ that appears in Theorem 4.4 is simply the worst-case scenario across supply and demand parameters of the social welfare ratio: $\Phi_1(n) = \min_{z \geq 0} \Phi_1(n, z)$. The bound of 99% is reached for $n = 3$ and $z = -n + 2\ \sqrt{n^2 - n}$.

Turning now to consumer surplus,

**Theorem 4.5.** *With $n$ symmetric generators ($n \geq 3$) competing through a supply function bidding mechanism, consumers lose at most 32% of their surplus compared to the centrally regulated solution.*

$$\frac{CS(Dereg.)}{CS(Centr.)} \geq \Phi_2(n) \geq 80\% \qquad \text{for } n \geq 3$$

*Proof.* See Appendix C.1. □

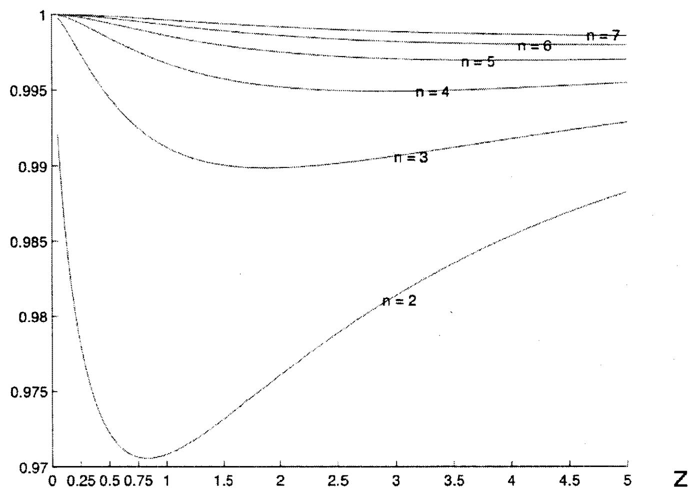


Figure 4-5: Welfare ratios for $n$ symmetric generators ($n$ ranging from 2 to 7) as functions of the unit-less cost parameter $z$.

Figure 4-6 depicts the social welfare ratio $\frac{CS(Dereg.)}{CS(Centr.)}$ as a function of the demand elasticity $z$ for oligopoly markets with various number of generators. We denote the curves of Figure 4-6 by $\Phi_2(n, z)$. Again, these functions are increasing with the number of generators uniformly across the market parameter $z$. The adverse effect of deregulation on consumers is less significant when the market has a large number of generators competing. The functions $\Phi_2(n, z)$ are also unimodal and surprisingly enough, they reach their minimum at the same point as the minimum of the social welfare ratios: $z = -n + 2\sqrt{n^2 - n}$. The bound $\Phi_2(n)$ that appears in Theorem 4.5 is simply the worst-case scenario across supply and demand parameters of the social welfare ratio: $\Phi_2(n) = \min_{z \geq 0} \Phi_2(n, z)$. The bound of 80% is reached for $n = 3$ and $z = -n + 2\sqrt{n^2 - n}$. The bounds presented in Section 4.3 are all tight since they are derived from closed form expressions of the profit, surplus and welfare ratios.

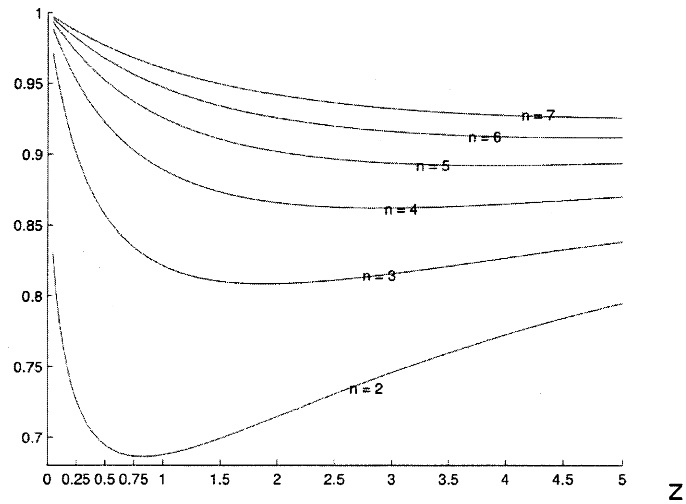


Figure 4-6: Consumer surplus ratios for $n$ symmetric generators ($n$ ranging from 2 to 7)
as functions of the unit-less cost parameter $z$.

## 4.4 Unconstrained asymmetric generators

### 4.4.1 The model

In this section, we assume that generators are heterogeneous, that is, they have different production costs. Since electricity is a homogeneous good ($D(p, \epsilon) = -m\, p + \epsilon$), the generators with the cheapest production costs dominate the market. Since they can produce cheaply, they can not only undercut competition and sell more electricity but they can also exercise a lot of market power by pricing their electricity up to the level of the more expensive generators. In describing an asymmetric market, it is no longer enough to simply count the number of generators, as these generators might be very different from each other. If, for example, one generator has production costs that are way more expensive than its competitors, it will probably not be competitive and it will not produce any electricity. In such a case, this generator is not significant and it should not be counted among the competing generators. As described in Section 4.2, the generators face quadratic production costs $C_i(q_i) = b_i q_i + 1/2 c_i q_i^2,\ i = 1, \ldots, n$.

The generators place bids to the system operator by submitting affine supply functions: $q_i(p) = \beta_i(p - \alpha_i)$, $i = 1, \ldots, n$. For every possible realization of the demand uncertainty $\epsilon$, generator $i$ attempts to maximize its profit (taking the supply bids of the competitors as given) by selecting $p_\epsilon$ and $q_{i.\epsilon}$ such that :

$$\forall \epsilon \geq 0, \qquad p_\epsilon,\ q_{i,\epsilon} = \underset{p,\ q_i}{\text{argmax}} \quad p\, q_i - C_i(q_i)$$
$$\text{s.t.} \quad q_i + \textstyle\sum_{j \neq i} q_j(p) = D(p, \epsilon),$$

Hence, for every realization of $\epsilon$, generator $i$ chooses a point $(p^*_\epsilon, q^*_{i,\epsilon})$ that maximizes the objective $p[D(p, \epsilon) - \sum_{j \neq i} q_j(p)] - C_i\left(D(p, \epsilon) - \sum_{j \neq i} q_j(p)\right)$ (since $q_i$ is the residual demand $q_i = D(p, \epsilon) - \sum_{j \neq i} q_j(p)$). Setting the derivative (with respect to $p$) to zero leads to :

$$q_{i.\epsilon} = [p_\epsilon - C'_i(q_{i,\epsilon})] \left( -\frac{dD}{dp} + \sum_{j \neq i} \frac{dq_j}{dp} \right) \tag{4.8}$$

Generator $i$ chooses the pair $(\alpha_i, \beta_i)$ of its supply bid $S_i(p) = \beta_i(p - \alpha_i)$ before knowing the realization of demand. The bid forces generator $i$ to produce $q_{i,\epsilon} = \beta_i(p_\epsilon - \alpha_i)$ while optimality condition (4.8) must be satisfied for all realizations of $\epsilon$. The two conditions can be combined into:

$$\beta_i(p_\epsilon - \alpha_i) = [p_\epsilon - b_i - c_i\beta_i(p_\epsilon - \alpha_i)] \left( m + \sum_{j \neq i} \beta_j \right) \qquad \forall \epsilon \geq 0 \tag{4.9}$$

Since supply must match demand $\sum_{j=1}^n \beta_j\, p_\epsilon - \sum_{j=1}^n \beta_j \alpha_j = -mp_\epsilon + \epsilon$, it is clear that the equilibrium price $p_\epsilon$ varies with uncertainty $\epsilon$. The only way optimality condition (4.9) can be satisfied $\forall \epsilon$, is that generator $i$ chooses the pair $(\alpha_i, \beta_i)$ so that the polynomials (in terms of $p_\epsilon$) on both sides of equation (4.9) are equal. We must have:

$$\beta_i = (1 - c_i\beta_i)(m + \sum_{j \neq i} \beta_j), \qquad \forall i \tag{4.10}$$

and

$$\alpha_i \beta_i = (b_i - c_i \beta_i \alpha_i)(m + \sum_{j \neq i} \beta_j), \qquad \forall i \tag{4.11}$$

Multiplying equation (4.10) by $\alpha_i$ and subtracting equation (4.11), we obtain:

$$(\alpha_i - \alpha_i \beta_i c_i)(m + \sum_{j \neq i} \beta_j) = (b_i - \alpha_i \beta_i c_i)(m + \sum_{j \neq i} \beta_j), \qquad \forall i$$

$$\Rightarrow \alpha_i = b_i, \qquad \forall i$$

**Proposition 4.2.** *The equilibrium for supply bids functions $S_i(p) = \beta_i(p - \alpha_i)$ is characterized by the system of equations:*

$$\begin{cases} \alpha_i = b_i \\ \beta_i = (1 - c_i \beta_i)(m + \sum_{j \neq i} \beta_j) \end{cases} \qquad i = 1, \ldots, n$$

*This system of equations has a unique solution vector $(\boldsymbol{\alpha}, \boldsymbol{\beta})$.*

*Proof.* We just derived the system of equations before the proposition. See [116] for a proof of existence and uniqueness. □

### 4.4.2 Market insights

In the remainder of Section 4.4, the goal is to derive some market insights and to evaluate the impact of deregulation on profit, consumer surplus and welfare ratios when suppliers are asymmetric. The task is complicated though due to the fact that the system of quadratic equations $\beta_i = (1 - c_i \beta_i)(m + \sum_{j \neq i} \beta_j)$, $i = 1, \ldots, n$, described in Proposition 4.2 does not lead to a closed form solution of the $\beta_i$'s as functions of the parameter $m$ of the demand and the costs parameters $c_i$'s. Unfortunately, there is no simple parametric solution $\beta_i = \Psi_i(m, c_1, \ldots, c_n)$. This makes calculating and bounding the performance ratios very difficult. To overcome this, we first reformulate the problem and establish bounds on the bid parameters $\beta_i$ themselves before investigating the performance ratios. Introducing the unit-less bid parameters $\widetilde{\beta}_i \triangleq \frac{\beta_i}{m}$, we remark though that the system of equations defining the $\widetilde{\beta}_i$'s only depends on $m$ and $c_i$'s through the unit-less parameters $z_i = mc_i$. That is, the following system of equations emerges:

$$\widetilde{\beta}_i = (1 - z_i\widetilde{\beta}_i)(1 + \sum_{j \neq i} \widetilde{\beta}_j), \; i = 1, \ldots, n \tag{4.12}$$

This is a unit-less reformulation of the supply function equilibrium. Since $S_i(p) = \beta_i(p - \alpha_i)$, generator $i$ offers less electricity for a given price when $\beta_i$ decreases. Alternatively, generator $i$ charges more when $\beta_i$ decreases. Equation (4.12) tells us that when its competitors charge more for electricity ($\sum_{j \neq i} \beta_j$ decreases), generator $i$ also charges more ($\beta_i$ decreases). Similarly, when its production cost $c_i$ (and hence $z_i$) increases, generator $i$ increases its electricity price ($\beta_i$ decreases).

In fact, an even stronger relationship exists between supply bids and marginal costs. The marginal cost of generator $i$ is: $C_i'(q_i) = b_i + c_i q_i$. On the other hand, the supply function provided by the generator is equivalent to a marginal price of: $p_i(q_i) = \frac{q_i}{\beta_i} + \alpha_i$ charged per unit of production. Even without solving for the $\widetilde{\beta}_i$'s, we can already establish that no generator will bid at a loss in this competitive setting.

**Theorem 4.6.** *Under free competition (supply function equilibrium), generators always bid a marginal price above their marginal production cost.*

*Proof.* Since $\alpha_i = b_i$, it is enough to prove that $\frac{1}{\beta_i} \geq c_i$ or equivalently $\widetilde{\beta}_i \leq \frac{1}{z_i}$. This is shown in Appendix C.2. □

In the rest of the section, we consider the case of generators facing purely quadratic production costs, that is $C_i(q_i) = 1/2 c_i q_i^2$ (i.e. $b_i = 0$ and hence $\alpha_i = b_i = 0$ from Proposition 4.2). Given the supply functions submitted by the generators, the equilibrium price matches supply with demand: $\sum_{i=1}^n \beta_i \, p = -mp + \epsilon$. This leads to the equilibrium price:

$$p_\epsilon^* = \frac{\epsilon/m}{1 + \sum_{i=1}^n \widetilde{\beta}_i} \tag{4.13}$$

The consumer surplus and the generators' profit can be expressed as (see Appendix C.2):

$$CS(Dereg.) = \frac{\epsilon^2}{2m}\left[1 - \frac{1}{1+\sum_{i=1}^{n}\widetilde{\beta}_i}\right]^2$$

$$\text{Profit}(Dereg.) = \frac{\epsilon^2}{m}\frac{1}{(1+\sum_{i=1}^{n}\widetilde{\beta}_i)^2}\left[\sum_{i=1}^{n}\widetilde{\beta}_i(1-\frac{z_i}{2}\widetilde{\beta}_i)\right]$$

We compare the deregulated setting to a centrally coordinated setting, where the system operator decides and dispatches the production quantities of the generators in order to maximize social welfare. In the absence of production constraints the system operator will dispatch the generators so that the equilibrium price paid to the generators exactly equals their marginal cost.

As discussed earlier, the marginal cost of generator $i$ is: $C_i'(q_i) = b_i + c_i q_i$. As a result, we can denote the inverse marginal cost by: $(C_i')^{-1}(p) = \frac{p-b_i}{c_i}$. The system operator must find a price such that generators produce at marginal costs and supply matches demand. Under the assumption $\alpha_i = b_i = 0$, this can be written as:

$$\sum_{i=1}^{n}(C_i')^{-1}(\widetilde{p}) = -m\widetilde{p} + \epsilon \qquad \Rightarrow \qquad \widetilde{p} = \frac{\epsilon/m}{1+\sum_{i=1}^{n}1/z_i}$$

Under the centrally coordinated scenario, the consumer surplus and the generators' profit are (as derived in Appendix C.2):

$$CS(Centr.) = \frac{\epsilon^2}{2m}\left[1 - \frac{1}{1+\sum_{i=1}^{n}\frac{1}{z_i}}\right]^2$$

$$\text{Profit}(Centr.) = \frac{\epsilon^2}{2m}\frac{1}{(1+\sum_{i=1}^{n}\frac{1}{z_i})^2}\left[\sum_{i=1}^{n}\frac{1}{z_i}\right]$$

As expected, free competition allows the generators to exercise some market power. They produce less electricity overall and charge a higher price.

**Theorem 4.7.** *Electricity is more expensive under free competition than under the* *centrally coordinated scenario, i.e. $p(Dereg.) \geq p(Centr.)$. As a result, consumers consume less electricity overall, $q(Dereg.) \leq q(Centr.)$*

*Proof.* As is established in Lemma 4.1 below, $\widetilde{\beta}_i \leq \frac{1}{z_i}$. This implies $p^* \geq \widetilde{p}$ and since $q = -m\, p + \epsilon$, we also have $q^* \leq \widetilde{q}$. $\square$

As in the previous section, our goal is to study the ratios $\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)}$, $\frac{CS(Dereg.)}{CS(Centr.)}$, and $\frac{TS(Dereg.)}{TS(Centr.)}$.

**Theorem 4.8.** *The ratios of generators profit, consumer surplus and social welfare under free competition compared to the centrally coordinated scenario $\left(\frac{Profit(Dereg.)}{Profit(Centr.)}, \frac{CS(Dereg.)}{CS(Centr.)}, \frac{TS(Dereg.)}{TS(Centr.)}\right)$ are independent of the demand stochasticity $\epsilon$. Moreover, they depend on the input parameters $m$ and $c_i$ only through the unit-less variables $z_i = mc_i$.*

*Proof.* We established in system (4.12) that the $\widetilde{\beta}_i$'s only depend on $m$ and $c_i$'s through the $z_i$'s. In the ratios below, only $\widetilde{\beta}_i$'s and $z_i$'s appear, thus proving the theorem.

$$\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)} = \frac{\frac{2}{(1+\sum_{i=1}^n \widetilde{\beta}_i)^2}\left[\sum_{i=1}^n \widetilde{\beta}_i(1-\frac{z_i}{2}\widetilde{\beta}_i)\right]}{\frac{1}{(1+\sum_{i=1}^n \frac{1}{z_i})^2}\left[\sum_{i=1}^n \frac{1}{z_i}\right]}$$

$$\frac{CS(Dereg.)}{CS(Centr.)} = \frac{\left[1-\frac{1}{1+\sum_{i=1}^n \widetilde{\beta}_i}\right]^2}{\left[1-\frac{1}{1+\sum_{i=1}^n \frac{1}{z_i}}\right]^2}$$

$$\begin{aligned}\frac{TS(Dereg.)}{TS(Centr.)} &= \frac{\frac{1}{(1+\sum_{i=1}^n \widetilde{\beta}_i)^2}\left[(\sum_{i=1}^n \widetilde{\beta}_i)^2 + (\sum_{i=1}^n \widetilde{\beta}_i) + (\sum_{i=1}^n \widetilde{\beta}_i(1-z_i\widetilde{\beta}_i))\right]}{\frac{1}{(1+\sum_{i=1}^n 1/z_i)^2}\left[(\sum_{i=1}^n 1/z_i)^2 + (\sum_{i=1}^n 1/z_i)\right]} \\ &= \frac{1-\frac{1}{1+\sum_{i=1}^n \widetilde{\beta}_i}}{1-\frac{1}{1+\sum_{i=1}^n 1/z_i}} + \frac{\left[\sum_{i=1}^n \widetilde{\beta}_i(1-z_i\widetilde{\beta}_i)\right] / \left[1+\sum_{i=1}^n \widetilde{\beta}_i\right]^2}{1-\frac{1}{1+\sum_{i=1}^n 1/z_i}}\end{aligned}$$

$\square$

In the remainder of the section, we approximate the performance ratios above with the goal to understand how the deregulation process affects consumers and generators depending on the market structure (the number of generators and their production costs). Since the $\widetilde{\beta}_i$'s cannot be obtained in closed form (as function of $z_i$'s), the first step is to bound them.

**Lemma 4.1.** *The slopes of the supply function bids $\widetilde{\beta}_i$'s can be bounded above and below with increasing degrees of accuracy. In particular, the following two bounds hold:*

$$\frac{1}{1+z_i} \leq \widetilde{\beta}_i \leq \frac{1}{z_i} \quad \textit{and} \quad \frac{1+\sum_{j\neq i}\frac{1}{1+z_j}}{1+z_i(1+\sum_{j\neq i}\frac{1}{1+z_j})} \leq \widetilde{\beta}_i \leq \frac{1+\sum j\neq i\frac{1}{z_j}}{1+z_i\sum_{j\neq i}\frac{1}{z_j})}$$

*Proof.* See Appendix C.2 for proof. □

A useful notion, we now introduce is the *equivalent number of generators* measure $r = \sum_i \frac{z_{min}}{z_i}$. $z_{min}$ is the unit-less (quadratic) production coefficient of the cheapest generator denoted $i_{min}$. The cheapest a generator is, the more pricing power it has and the more energy it will provide to the market. When the other generators have a much higher production coefficient than $z_{min}$, $r = 1 + \sum_{j\neq i_{min}} \frac{z_{min}}{z_j} \to 1$, pointing to the fact that there is only 1 significant player. In contrast, when all the generators have the same production coefficient $z_{min}$, $r \to n$ as there are $n$ significant generators. We can now use $r$ to approximate the consumer surplus ratio.

**Theorem 4.9.** *The ratio of consumer surplus under free competition compared to the centrally coordinated scenario can be lower and upper bounded by simple functions of the $z_i$'s:*

$$\left[\frac{1-\frac{1}{1+\sum_{i=1}^n\frac{1}{1+z_i}}}{1-\frac{1}{1+\sum_{i=1}^n\frac{1}{z_i}}}\right]^2 \leq \frac{CS(Dereg.)}{CS(Centr.)} \leq \left[\frac{1-\frac{1}{1+\sum_{i=1}^n\frac{1}{z_i}-\frac{1}{1+\sum_{i=1}^n\frac{1}{z_i}}\left(\sum_{i=1}^n\frac{1}{z_i^2}\right)}}{1-\frac{1}{1+\sum_{i=1}^n\frac{1}{z_i}}}\right]^2 \leq 1 \tag{4.14}$$

*This ratio can be further lower bounded by a function of the minimum production cost of the generators $z_{min}$ and a measure of the equivalent number of generators $r$:*

$$\frac{CS(Dereg.)}{CS(Centr.)} \geq \left[\frac{1 - \frac{1}{1+\frac{r}{1+z_{min}}}}{1 - \frac{1}{1+\frac{r}{z_{min}}}}\right]^2 = \left(1 - \frac{1}{1+r+z_{min}}\right)^2 \qquad (4.15)$$

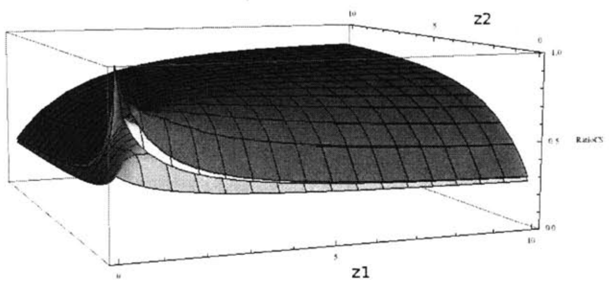


Figure 4-7: Consumer surplus ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds.

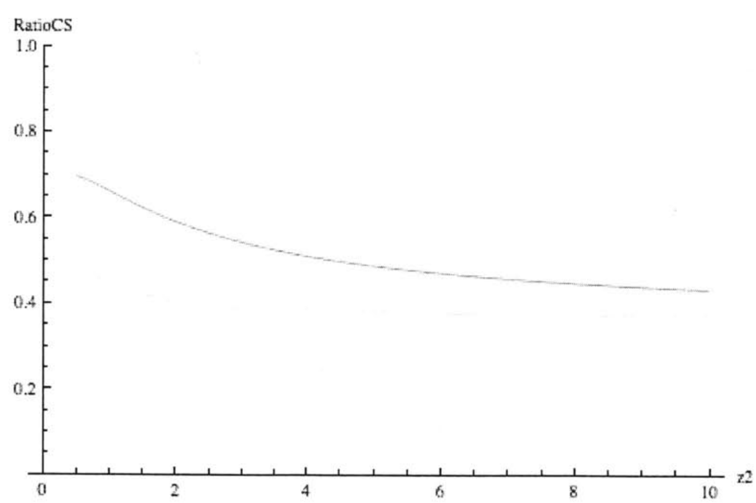


Figure 4-8: Consumer surplus ratio for a duopoly ($z_1 = 0.5$, $z_2$ varies) with simplified lower bound.

*Proof.* Bound (4.14) is directly derived from the upper and lower bounds on $\widetilde{\beta}_i$ as the funciton $1 - \frac{1}{1+\sum_{i=1}^n \widetilde{\beta}_i}$ is increasing with $\widetilde{\beta}_i$'s. To prove bound (4.15), we observe that

$1 \leq \frac{z_i}{z_{min}}$. Plugging this inequality into the lower bound (4.14) gives rise to:

$$\frac{CS(Dereg.)}{CS(Centr.)} \geq \left[\frac{1 - \frac{1}{1+\sum_{i=1}^{n} \frac{1}{z_i/z_{min}+z_i}}}{1 - \frac{1}{1+\sum_{i=1}^{n} \frac{1}{z_i}}}\right]^2$$

$$= \left[\frac{1 - \frac{1}{1+\frac{r}{1+z_{min}}}}{1 - \frac{1}{1+\frac{r}{z_{min}}}}\right]^2$$

$\square$

Figure 4-7 illustrates the consumer surplus bounds in the case of a duopoly market. The surfaces represent the consumer surplus ratio as a function of the market parameters $(z_1, z_2)$. The middle surface is the actual consumer surplus ratio, while the top and bottom surfaces are the upper and lower bounds of (4.14) respectively. Notice the proximity of the bounds to the exact consumer surplus ratio.

It is also worth noting that the ratio $\left[\left(1 - \frac{1}{1+\frac{1}{1+z_1}}\right) / \left(1 - \frac{1}{1+\frac{1}{z_1}}\right)\right]^2$ is the consumer surplus ratio of a monopolistic generator with quadratic coefficient $z_1$ (see Section 4.3.1). As a result, bound (4.15) above is the monopolistic bound modified to account for the number of significant generators through measure $r$.

Bound (4.15) is illustrated in Figure 4-8 for a duopoly market. The top curve is the consumer surplus ratio, the bottom curve is the bound. The bound is increasing with the measure $r$ of equivalent number of generators and with the lowest production coefficient $z_{min}$. In essence, the deregulation process becomes less and less costly to the consumers as either (i) there are more generators in the market or (ii) as the production costs of the generators increase. The reason is that in both cases, the generators cannot exercise a lot of market power. When there are many generators, competition prevents the exercise of extreme market power. Similarly, when production costs are expensive, the generators don't have much room left to distort price beyond the already expensive centrally coordinated price.

For example, when there are at least three equivalent generators ($r \geq 3$) and $z_{min} \geq 1$, the loss of consumer surplus due to deregulation cannot exceed 36%.

Let us now focus on the analysis of the profit ratio.

**Theorem 4.10.** *The generators are always better off in terms of total profit under free competition than under central coordination. Moreover the ratio of generators' profit under free competition, compared to the centrally coordinated setting, can be upper bounded by a function of the $z_i$'s. That is,*

$$1 \leq \frac{\mathit{Profit(Dereg.)}}{\mathit{Profit(Centr.)}} \leq \frac{\frac{1}{(1+\sum_{i=1}^{n}\frac{1}{1+z_i})^2}\sum_{i=1}^{n}\frac{2+z_i}{(1+z_i)^2}}{\frac{1}{(1+\sum_{i=1}^{n}\frac{1}{z_i})^2}\left[\sum_{i=1}^{n}\frac{1}{z_i}\right]} \tag{4.16}$$

*The ratio of generators' profit can be further upper bounded by a function of the minimum production cost of the generators $z_{min}$ and a measure of the equivalent number of generators $r$. That is,*

$$\frac{\mathit{Profit(Dereg.)}}{\mathit{Profit(Centr.)}} \leq \frac{(2+z_{min})(r+z_{min})^2}{z_{min}(1+r+z_{min})^2} \tag{4.17}$$

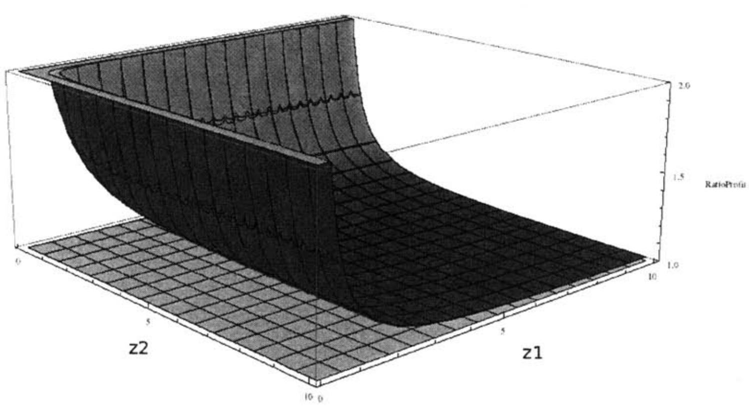


Figure 4-9: Profit ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds (4.16).

*Proof.* See Appendix C.2 for proof. □

Figure 4-9 illustrates the profit ratio for a duopoly market as a function of the market parameters $z_1, z_2$. The middle surface is the actual profit ratio, while the top surface is the upper bound of (4.16) and the bottom surface is simply the lower bound

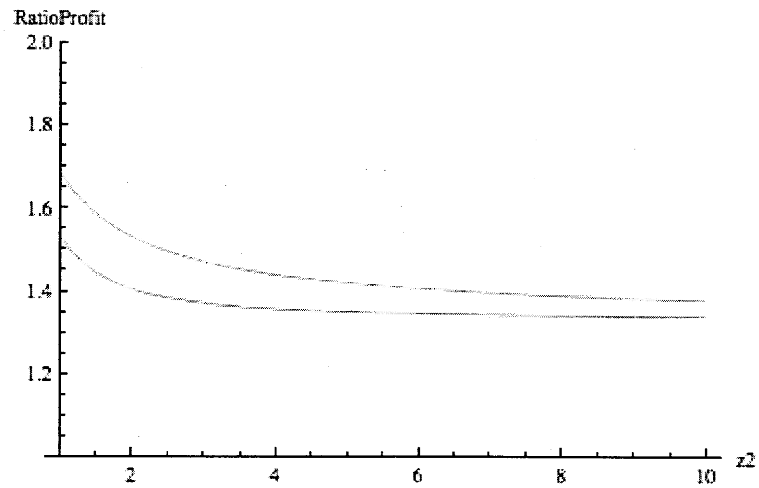


Figure 4-10: Profit ratio for a duopoly ($z_1 = 1$, $z_2$ varies) with simplified upper bound (4.17).

1. It is worth noticing the quality (i.e. the proximity) of our upper bound to the actual profit ratio.

When $r = 1$, bound (4.17) reaches $\frac{(1+z_{min})^2}{z_{min}(2+z_{miin})}$ which is the profit ratio of a monopolistic generator with quadratic coefficient $z_{min}$ (see Section 4.3.1). As a result, bound (4.17) can be viewed as the monopolistic bound modified to account for the number of equivalent generators $r$.

Bound (4.17) is illustrated in Figure 4-10. The bottom curve is the actual profit ratio, while the top curve is the upper bound. The bound decreases with $z_{min}$, but increases with $r$. The intuition is that when $z_{min}$ increases, the generators can exercise less market power so that the profit gained under deregulation decreases. When the number of significant generators $r$ increases on the other hand, even though each individual generator can exercise less market power, there are more generators increasing their profit through deregulation. The global gain in generators' profit from deregulation increases.

For example, when there are less than five significant generators ($r \leq 5$) and $z_{min} \geq 5$, generators taking advantage of deregulation can increase their profit by at most 16%.

We finally analyze the loss of welfare for society as a whole (generators and consumers). Again, we can bound it by two simple functions of the $z_i$'s:

**Theorem 4.11.** *The ratio of social welfare under free competition compared to the social welfare under central coordination can be upper and lower bounded by:*

$$\frac{\left[\frac{\sum_{i=1}^{n} 1/(1+z_i)}{1+\sum_{i=1}^{n} 1/(1+z_i)}\right]^2 + \frac{\sum_{i=1}^{n} 1/z_i}{(1+\sum_{i=1}^{n} 1/z_i)^2}}{1 - \frac{1}{1+\sum_{i=1}^{n} 1/z_i}} \leq \frac{TS(Dereg.)}{TS(Centr.)} \leq 1 \tag{4.18}$$

*The ratio of social welfare can be further lower bounded by a function of the minimum production cost of the generators $z_{min}$ and the measure $r$ of equivalent number of generators. That is,*

$$\frac{z_{min}}{r + z_{min}} + \frac{r\ (r + z_{min})}{(1 + r + z_{min})^2} \leq \frac{TS(Dereg.)}{TS(Centr.)} \leq 1 \tag{4.19}$$

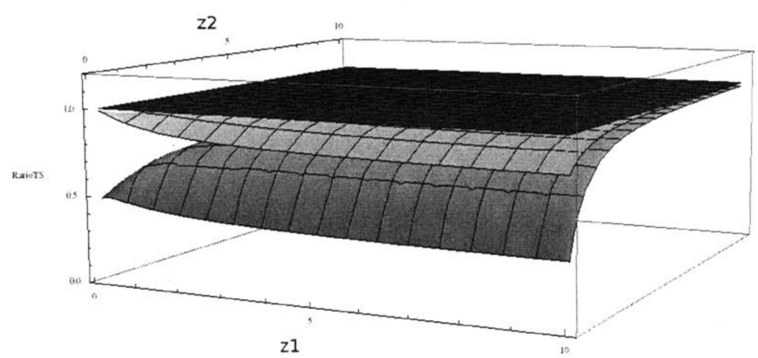


Figure 4-11: Social welfare ratio for a duopoly (as function of $z_1, z_2$) with upper and lower bounds.

*Proof.* See Appendix C.2 for proof. □

Figure 4-11 illustrates the social welfare ratio for a duopoly market. The middle surface is the actual welfare ratio, the bottom surface is the lower bound of (4.18) and the top surface is the upper bound 1. Bound (4.19) is illustrated in Figure 4-12. It increases with $z_{min}$. For example, when the quadratic cost coefficient of the

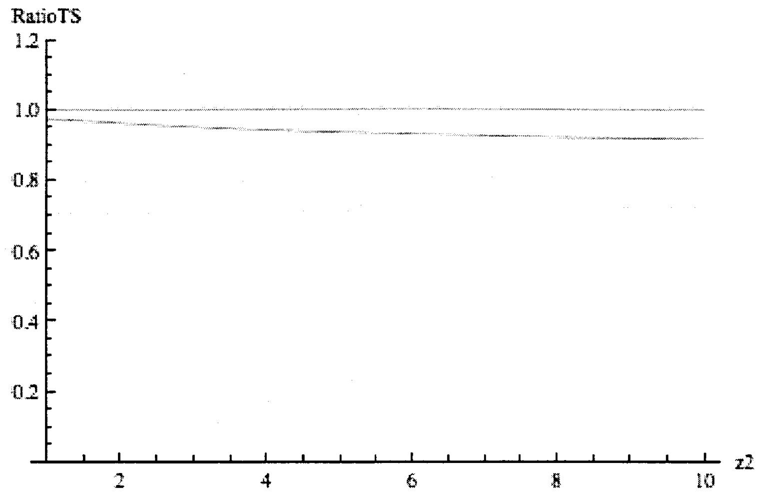


Figure 4-12: Social welfare ratio for a duopoly ($z_1 = 1$, $z_2$ varies) with simplified upper and lower bounds.

generators is above 1 (i.e. $z_{min} \geq 1$), the loss of social welfare cannot exceed 30% and when $z_{min} \geq 5$, the loss is bounded by 10%.

## 4.5 Performance under general network constraints

In this section, we test the supply function bidding mechanism by simulating it on two electricity systems: the classical IEEE 30-Bus system [107] and a 5-bus system modeling a subnetwork of the New England electricity grid. The challenge here is to model competition between electricity generators in the presence of the physical dispatch constraints of an electricity network. Even without the presence of physical constraints, modeling supply function equilibria is challenging as shown in [75]. There are only a handful of papers ([16], [64]) analyzing supply function competition in the presence of constraints. While these papers are computationally oriented, our focus in this chapter is on evaluating the performance of deregulation for the market participants. We begin by describing the technical challenges that arise in simulating our model and explain how we resolve them. We then discuss the performance of deregulation in the 30-bus and the 5-bus systems. We compare the performance ratios obtained on these systems to the bounds established in the previous sections.

### 4.5.1 Simulation method

As described in the introduction, we model the electricity market as a bi-level (Stackelberg) game. The generators attempt to maximize their profit by submitting a supply function to the system operator who in turn attempts to maximize the welfare of society. The generators are the leaders of this Stackelberg game, the system operator is the follower. Each generator solves a best response problem by choosing its supply bid to maximize its profit taking the bids of the other generators as given and anticipating the dispatch of the system operator. After collecting the bids of the generators, the system operator indeed decides the electricity dispatch to maximize social welfare while satisfying the electricity network constraints. We use the traditional DC linear approximation of electricity flows to model network constraints. See [74] for references on the subject and a detailed description.

The optimization of the system operator can be represented as:

$$
\begin{aligned}
(p^*, \mathbf{q}^*) \quad = \quad & \underset{\mathbf{q},p}{\text{argmax}} \quad ProxyTS(p, \mathbf{q}) \\
& \text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} q_i = D(p) & \text{(Supply = Demand)} \\ 0 \leq q_i \leq \beta_i(p - \alpha_i) & \text{(Supply bids)} \\ \Phi q \leq \Delta & \text{(Network constraints)} \end{cases}
\end{aligned}
\tag{4.20}
$$

Notice that the optimal solution $(p^*, \mathbf{q}^*)$ depends on the supply bids $(\alpha_i, \beta_i)$. Under our modeling assumptions of Section 4.2, there is a map that gives rise to a unique dispatch $(p^*, \mathbf{q}^*)$ for each set of bids $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ (see [69] for proof). We will denote this mapping by the optimal dispatch function OPTIDISP: $(p^*, \mathbf{q}^*) = \text{OPTIDISP}(\boldsymbol{\alpha}, \boldsymbol{\beta})$.

Anticipating the optimization of the system operator, each generator $i$ maximizes its profit by solving:

$$
\begin{aligned}
\max_{\alpha_i, \beta_i} \quad & p^* \; q_i^* - C_i(q_i^*) \\
\text{s.t.} \quad & (\mathbf{q}^*, p^*) = \text{OPTIDISP}(\alpha_i, \beta_i, \boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}_{-i})
\end{aligned}
\tag{4.21}
$$

, assuming the bids $(\boldsymbol{\alpha}_{-i}, \boldsymbol{\beta}_{-i})$ of the other generators as fixed.

In order to simulate the game played between the generators, the first step is to be able to simulate the optimal response of one of the generators to the bids of the other generators. This alone is a difficult problem because it is a bi-level optimization: the constraint of optimization (4.21) is itself the result of an optimization problem. The entire game between the generators is therefore a bi-level game.

Our first step toward simulating this game is to reformulate the optimization problem (4.20) of the system operator. Due to the concavity of the problem, the problem is equivalent to its KKT conditions. We can represent the equations and complementarity conditions of the KKT system through a mixed complementarity problem (MCP) (see [42], [43], [27] for references):

$$z = \begin{pmatrix} p^* \\ \mathbf{q} \end{pmatrix} \geq 0 \quad \perp \quad F_{\alpha_i,\beta_i,\boldsymbol{\alpha}_{-i},\boldsymbol{\beta}_{-i}}(z) \geq 0 \tag{4.22}$$

We do not detail $F_{\alpha_i,\beta_i,\boldsymbol{\alpha}_{-i},\boldsymbol{\beta}_{-i}}(z)$ explicitly here but it is derived from the constraints and the KKT optimality conditions of problem (4.20) in the standard way described in [42]. Problem (4.22) is equivalent to problem (4.20) in the sense that the solutions to the two problems are the same.

The bi-level optimization of generator $i$ can now be reformulated into a single level optimization problem:

$$\begin{aligned} \max_{\alpha_i,\beta_i} \quad & p^* \, q_i^* - C_i(q_i^*) \\ \text{s.t.} \quad & \left\{ \begin{pmatrix} p^* \\ \mathbf{q} \end{pmatrix} \geq 0 \quad \perp \quad F_{\alpha_i,\beta_i,\boldsymbol{\alpha}_{-i},\boldsymbol{\beta}_{-i}} \begin{pmatrix} p^* \\ \mathbf{q} \end{pmatrix} \geq 0 \right. \end{aligned} \tag{4.23}$$

Nevertheless, this problem is still hard to solve, even though it has a single level, since the constraints include complementarity conditions which are combinatorial in nature. For this reason, this class of problems is called an MPEC (maximization problem with equilibrium constraints) which in general is NP-hard to solve (see [83]).

Going back to the equilibrium of the overall game, each generator needs to solve this MPEC. We are looking for a quadruple $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{q}, p^*)$ which satisfies optimality in problem (4.23) for all generators $i = 1, \dots, n$ simultaneously. This in turn is an equilibrium problem with equilibrium constraints (EPEC) (see [16] for reference).

In order to simulate competition (the simultaneous optimization of the generators), we use NIRA, a powerful tool to compute Nash equilibria in a constrained multiplayer game. NIRA [79] was developed by Jacek Krawczyk and James Zuccollo and combines the use of the Nikaido Isoda function and a relaxation algorithm. At a high level, the algorithm is based on the ability to compute the best response of one player to the bids of the other players and to iterate the process across players a number of times in a particular fashion until convergence is reached and no player wants to modify its bid. See [79] for further explanations.

Compared to NIRA though, our problem has an additional level of complexity. NIRA is well suited to find equilibria of Nash games with simple coupling constraints (equalities or inequalities) tying the players. However, the constraint tying the generators together in our game are complementarity constraints which are discrete in nature. In order to converge, NIRA needs to be able to solve the best response problem (4.23) of a given generator quickly. After testing several methods to approximate MPECs, we decided to use KNITRO [142]. KNITRO is a fast numerical solver for large scale non-linear optimization problems. It offers an efficient MPEC solver that we use to compute the generators' best responses.

We combine the use of NIRA and KNITRO in a MATLAB setting. NIRA is directly available in MATLAB but KNITRO isn't. We call the KNITRO solver via Tomlab [67]. We ran simulations using a Windows XP SP3 machine with an Intel Core Duo 3.27 Ghz and 512 MB of RAM. We use MATLAB 7.7.0 (R2008b) and NIRA version 3.

Let us now delve into the analysis of the simulations of the two electricity networks. We begin with the 30-bus system.

### 4.5.2 The 30-bus system

As mentioned, this example comes from the Power Systems Test Case Archive [107], which is a set of electricity network models provided by the University of Washington for the purpose of simulating and testing dispatch algorithms. The 30-bus system in particular, has been the subject of a few electricity market studies. Maiorano et al. [84] proposed a dynamic non-collusive Cournot model of electricity bidding and simulated it on the 30-bus system. Later, Contreras et al. [30] analyzed static Cournot equilibria in constrained electricity markets, also illustrating their model on the 30-bus system.

We use here the same set of assumptions as in Contreras et al. [30] in order to be able to compare the supply function bidding mechanism of this chapter to the Cournot model. We show that the supply function bidding achieves better social welfare than quantity (Cournot) bidding and that it is "almost" optimal in terms of social welfare.

Figure 4-13 shows a diagram of the 30-bus system. It is assumed that the system has six generators, split between three companies. The capacity of each generator and each company is provided below in Table 4.1: each generator can produce between $P_g^{min}$ and $P_g^{max}$ and each company supplies between $P_C^{min}$ and $P_C^{max}$.

| company # | generator # | $P_g^{min}$ [MW] | $P_g^{max}$ [MW] | $P_C^{min}$ [MW] | $P_C^{max}$ [MW] |
|---|---|---|---|---|---|
| 1 | 1 | 0 | 80 | 0 | 80 |
| 2 | 2 | 0 | 80 | 0 | 130 |
| | 3 | 0 | 50 | | |
| 3 | 4 | 0 | 55 | 0 | 125 |
| | 5 | 0 | 30 | | |
| | 6 | 0 | 40 | | |

Table 4.1: IEEE 30-bus system market data.

The cost for company $i$ to generate $q_i$ units is assumed quadratic of the form: $C_i(q_i) = b_i\ q_i + 1/2c_i\ q_i^2$. The production cost coefficients for each company are given in Table 4.2. Demand is a strictly decreasing function of the price $p$. Consistent with

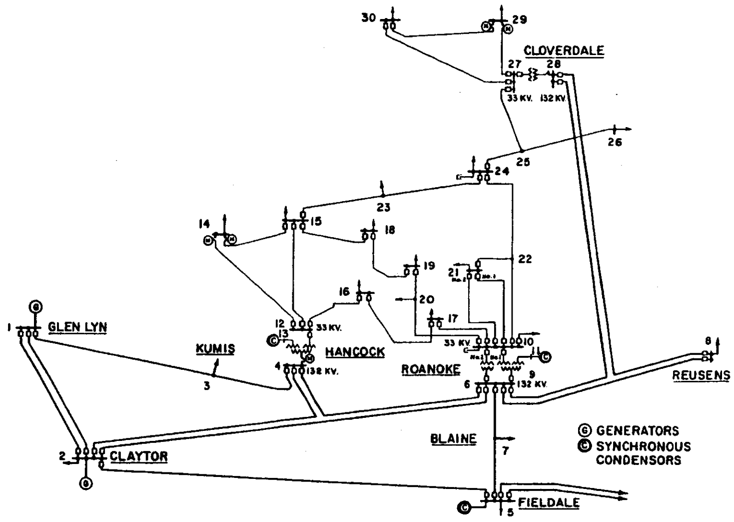


Figure 4-13: 30-bus power system test case.

the assumption in this chapter and in [30], we consider a linear demand function: $D(p) = 250 - 0.5p$.

| Company # | $c_i$ [$/MW²h] | $b_i$ [$/MWh] |
|---|---|---|
| #1 | 0.04 | 2.000 |
| #2 | 3.194*10-2 | 1.354 |
| #3 | 1.124*10-2 | 3.087 |

Table 4.2: 30-bus system generating cost data.

Each generator bids a linear supply function $S_i(p) = \beta_i(p - \alpha_i)$ to the system operator to maximize its profit. The system operator then decides the dispatch to maximize the proxy of social welfare (also denoted TS for total surplus) taking into account the capacity of the generating companies given in Table 4.1. The final bids, the production quantities and the profit of each company under our supply function equilibrium model are presented in Table 4.3. The social welfare achieved under the supply function bidding mechanism is $61,643.

We compare these results with the output of the Cournot model from [30]. Notice that compared to Contreras et al. [30], we increased the demand from $D(p) =$

| Company # | $\beta_i$ [$/MW²h] | $\alpha_i$ [$/MWh] | q [MW] | Profit [$] |
|---|---|---|---|---|
| #1 | 0.0219 | 7.2612 | 80.0000 | 1,038 |
| #2 | 0.1799 | 1.2648 | 85.1557 | 1,181 |
| #3 | 0.1585 | 4.4475 | 76.5529 | 1,000 |

Table 4.3: 30-bus system supply function equilibrium output.

$189.2 - 0.5p$ to $D(p) = 250 - 0.5p$. We did this, because with the original demand model, the capacity constraints of the generators are not active neither under the Cournot model nor under the supply function equilibrium (SFE) model. We do not show simulation results under the original demand model as this case is equivalent to the unconstrained case which we addressed in Sections 4.3 and 4.4. Under the increased demand, company 1 produces up to capacity (80 MW) under the SFE model, making the presence of constraints meaningful. We only made this change to make the problem more interesting. In order to compare our results with those of [30], we thus had to re-implement their algorithm and simulate it under this new demand. For consistency, we first checked that our implementation of their Cournot model produced the outcome reported in their paper under the original demand scenario. Below in Table 4.4, we present the outcome of the Cournot model under the increased demand scenario. The companies all produce less eletricity. They make more profit (approximately seven times more) but at the cost of reducing social welfare by 6% (from $61,643 to $57,851).

| Company # | q [MW] | Profit [$] |
|---|---|---|
| #1 | 61.7152 | 7,693 |
| #2 | 62.2780 | 7,819 |
| #3 | 62.0573 | 7,723 |

Table 4.4: 30-bus system Cournot output.

We conclude the comparison by computing the quantities of interest also under the centrally coordinated setting, where the system operator knows the generators costs and decides the dispatch to maximize social welfare. The outcome is described in Table 4.5. The maximum achievable social welfare (reached under this setting) is \$61,697. Under this setting however, generators make almost no profit.

| Company # | q [MW] | Profit [\$] |
|---|---|---|
| #1 | 55.8187 | 62.3146 |
| #2 | 90.1299 | 129.7307 |
| #3 | 101.9350 | 58.3960 |

Table 4.5: 30-bus system centrally coordinated output.

Our supply function equilibrium mechanism not only outperforms the Cournot model in terms of social welfare but it produces 99.9% of the maximum social welfare attained under the centralized scenario. The actual social welfare performance of the deregulated system on the 30-bus system is definitely better than the lower bound guarantee of equation (4.18) in Section 4.4. The bound only guarantees a welfare performance above 56% for 3 unconstrained asymmetric generators with production costs $z_i$'s as in Table 4.2. The actual performance of 99.9% is even better than the 99% guarantee of Theorem 4.4 for the case of 3 unconstrained symmetric generators. Moreover, the SFE mechanism allows the generating companies to make a decent profit of roughly \$1000 each out of a total welfare of \$61,697 instead of a small profit of roughly \$100. This could help answer the daunting problem states face to create a market process that encourages generators not only to produce electricity but also to invest in new capacity. Finally, implementing a centrally coordinated dispatch requires to monitor the production costs of the generators and this control comes at some cost, not represented in our model. If we account for these monitoring costs in the centrally coordinated setting (these costs reduce social welfare), the SFE mechanism will undoubtedly turn out to be more efficient for society than the centralized solution.

### 4.5.3 The 5-bus system

The 30-bus system is interesting to simulate because it has been studied by a number of other papers ([84], [30]). It allows us to compare the outcome of our supply function equilibrium model with the classical Cournot model as described in these papers. Yet, the representation of the 30-bus system in [30] does not include transmission constraints tying the generators together. It only models capacity constraints at the generators' level and at the companies' level.

In order to evaluate the full benefit of our supply function equilibrium model, we need to consider a power system containing realistic electrical grid coupling constraints. We use a 5-bus system modeling part of the New England electrical grid. It was provided to us courtesy of Andy Sun [1] who interned at the New England ISO. The goal in carrying simulations on the 5-bus system is to show that the superiority of the SFE model over the Cournot model and the quasi-optimality of the SFE model in terms of social welfare demonstrated on the 30-bus system still hold in this more complex/realistic setting.

A diagram of the 5-bus system is given in Figure 4-14. The system has five buses (nodes) denoted A to E and six electric lines (AB, AE, BC, CD, AD, DE). There are six generating companies supplying electricity to the network: Brighton, Alta, Park City, Solitude 1 & 2 and Sundance. The demand (load) is concentrated at a single bus, bus B. The diagram also shows the reactance $x_{ij}$ of all the electric lines. The reactance values are expressed in ohms and the % symbol does not refer to any ratio, simply to $10^{-2}$.

Using the DC approximation model, we can represent the network constraints with a linear system of inequalities: $A_{net}\, \mathbf{q} \leq d_{net}$. $\mathbf{q}$ simply denotes the vector of electricity generation (in MW) by the six generators. $A_{net}$ is called the distribution factor matrix and relates the electricity generated by the generators to the amount of electricity flowing on the lines. It can be calculated with the information given in Figure 4-14 (see [140] for details on the derivation). $d_{net}$ is the thermal capacity of

[1] PhD Student in the Operations Research Center at MIT

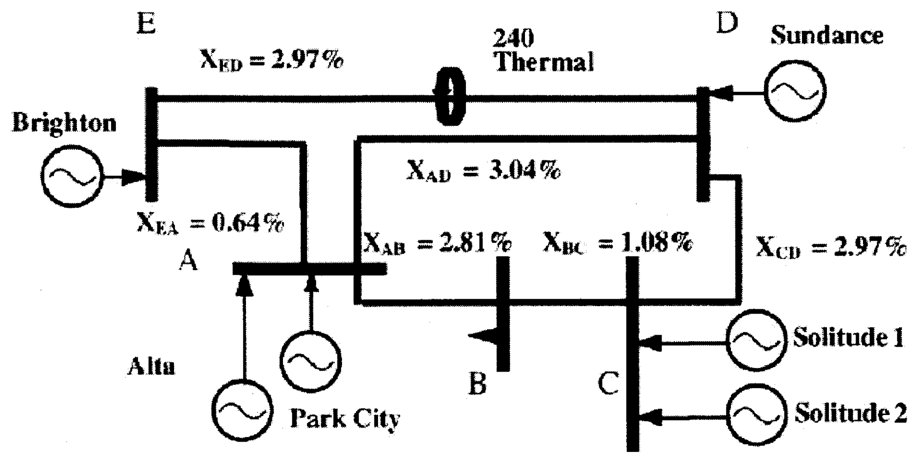


Figure 4-14: 5-bus power system diagram.

the six lines, the maximum electricity that can flow through them. $A_{net}$ and $d_{net}$ are presented in Table 4.6 & 4.7 respectively.

| A_net | Alta | Park | Brighton | S1 | S2 | Sundance |
|---|---|---|---|---|---|---|
| $F_{EA}$ | -0.0812 | -0.0812 | 0.7988 | 0.0008 | 0.0008 | 0.2419 |
| $F_{ED}$ | 0.0812 | 0.0812 | 0.2012 | -0.0008 | -0.0008 | -0.2419 |
| $F_{AB}$ | 0.8224 | 0.8224 | 0.7704 | 0.0018 | 0.0018 | 0.5293 |
| $F_{AD}$ | 0.0964 | 0.0964 | 0.0284 | -0.0009 | -0.0009 | -0.2873 |
| $F_{BC}$ | -0.1776 | -0.1776 | -0.2296 | -0.9982 | -0.9982 | -0.4707 |
| $F_{CD}$ | -0.1776 | -0.1776 | -0.2296 | 0.0018 | 0.0018 | -0.4707 |

Table 4.6: 5-bus distribution factor matrix.

| d_net | |
|---|---|
| $F_{EA}$ | 425 |
| $F_{ED}$ | 350 |
| $F_{AB}$ | 2000 |
| $F_{AD}$ | 300 |
| $F_{BC}$ | 400 |
| $F_{CD}$ | 400 |

Table 4.7: 5-bus line thermal capacities.

The generators have quadratic production costs: $C_i(q_i) = b_i\ q_i + 1/2c_i\ q_i^2$. The production cost coefficients for each company are given in Table 4.8. Demand is a strictly decreasing linear function of the price $p$: $D(p) = \epsilon - 0.01\ p$. As discussed in Sections 4.2 and 4.4, the SFE model is designed to handle intraday demand uncertainty. To account for this uncertainty, we model demand as a discrete random variable with three possible realizations: $\epsilon = [2936, 4404, 6606]$. The demand can take one of these three values with a probability of 1/3.

| Prod_cost | $c_i$ [\$/MW²h] | $b_i$ [\$/MWh] |
|---|---|---|
| Alta | 2.5 | 0 |
| Park | 3 | 1 |
| Brighton | 2 | 2 |
| S1 | 6 | 1 |
| S2 | 14 | 6 |
| Sundance | 5 | 1.6 |

Table 4.8: 5-bus system generating cost data.

The supply bid parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ of the SFE model are independent of the realization of demand uncertainty. This is due to the fact that the bids are chosen by the generators before demand realization. The production quantities, generators' profit and the welfare of society on the other hand, depend on the demand realization. For simplicity, we present the results for the demand realization: $\epsilon = 4404$. The insights derived hereafter hold true across all realizations though. The final bids, the production quantities and the profit of each company under the supply function equilibrium model we presented, can be found in Table 4.9. The total profit of all companies is \$64 Million and the social welfare is \$954 Million. The global SFE production quantity is 4217.55 MW.

| Company | $\beta_i$ [\$/MW²h] | $\alpha_i$ [\$/MWh] | q [MW] | Profit [\$] |
|---|---|---|---|---|
| Alta | 0 | 0 | 534.97 | 9616904 |
| Park | 0 | 0 | 534.97 | 9544822 |
| Brighton | 0 | 0 | 0 | 0 |
| S1 | 0 | 33.49 | 554.97 | 9423076 |
| S2 | 0 | 38.72 | 480.00 | 7334068 |
| Sundance | 1559.33 | 0 | 2112.65 | 28229609 |

Table 4.9: 5-bus supply function equilibrium dispatch.

We compare these results with the output of the Cournot model. Here generators simply bid generation quantities after the realization of the demand: $\epsilon = 4404$. We present in Table 4.10 the outcome of the Cournot model. Generators produce significantly less electricity overall at 3740.66 MW, that is 11% less than under SFE competition. They make a profit of $242 Million which is roughly 4 times the global profit achieved under SFE competition. However, this profit increase under Cournot competition comes at the cost of reducing the welfare of society by $12 Million or 1% compared to SFE competition (from $954 Million under SFE competition to $942 Million under Cournot competition). Under Cournot competition, generators still have the ability to exercise consequent market power.

| Company | q [MW] | Profit [$] |
|---|---|---|
| Alta | 934.88 | 60922000 |
| Park | 644.01 | 42097595 |
| Brighton | 520.74 | 34270937 |
| S1 | 558.41 | 36105814 |
| S2 | 518.85 | 32529838 |
| Sundance | 563.77 | 36601756 |

Table 4.10: 5-bus Cournot dispatch.

We conclude the comparison by computing the centrally coordinated scenario. The outcome is described in Table 4.11. The maximum achievable social welfare (reached under this scenario) is $961 Million. The overall profit of the generators under central coordination is small at $21 Million. Generators produce 4335.03 MW which is fairly close to the SFE production but 13% more than the Cournot production. Again, SFE competition is able to prevent excessive exercise of market power on the part of generators, something Cournot competition fails to achieve. Under SFE competition, generators make more profit than under centralized dispatch but a lot less than under Cournot competition. They produce almost as much electricity as under centralized dispatch and achieve 99.2% of the optimal social welfare attained under centralized dispatch. The social welfare of 99.2% is above the 73% guarantee of bound (4.18) in Section 4.4 for the case of 6 asymmetric unconstrained generators with parameters $z_i$'s as in Table 4.8. As mentioned in the previous section, our calculation of social welfare under centralized dispatch did not take into account the practical costs involved in gathering the production costs of the generators and in insuring compliance with the decided dispatch. If we also Integrate these costs in the centralized dispatch (these costs reduce social welfare) then, SFE appears to be an even more efficient mechanism to dispatch electricity.

| **Company** | **q [MW]** | **Profit [$]** |
|---|---|---|
| **Alta** | 842.74 | 4924592 |
| **Park** | 701.95 | 4101527 |
| **Brighton** | 485.77 | 3113418 |
| **S1** | 1147.32 | 3962884 |
| **S2** | 491.35 | 1695917 |
| **Sundance** | 665.91 | 3483114 |

Table 4.11: 5-bus centrally coordinated dispatch.

## 4.6 Conclusions

This chapter analyzes the costs and benefits of deregulating an electricity market. It introduces a new model of competition between generating firms. Based on supply function equilibrium concepts, our model provides a more accurate representation of actual electricity markets. It describes explicitly the role of the system operator as a maximizer of social welfare, and its bi-level formulation allows the modeling of electrical network constraints. After introducing the model, the chapter investigates the impact of deregulation on consumers, generators and society as a whole. The chapter begins with the case of symmetric generators first, ignoring electrical constraints. This simplified scenario highlights the dependence between the performance of a free market and the number of generators competing in the market. With a monopolistic generator, deregulation is not advisable: it can reduce the welfare of society by up to 25% and it can take away up to 75% of consumer surplus. Moving from a monopolistic to a duopolistic market drastically improves performance. While generators still benefit from the deregulation process, the loss of social welfare is reduced to 3% and the loss of consumer surplus decreases from 75% to 32%. As more generators enter the electricity market, the trend continues: the loss of social welfare and consumer surplus become smaller. The chapter then turns to the case of asymmetric generators. While the analysis is a lot more involved in this case, the basic insights are consistent with the symmetric case. Generators always benefit from the deregulation process. The amount of market power they can exercise (and the profit they can extract) decreases as more generators compete to supply electricity. On the other hand, consumers lose some surplus in the process of deregulation (as generators are able to drive prices up under free competition) but this loss decreases rapidly as more "significant" generators participate in the market. To quantify this phenomenon. this chapter defines a measure $r$ of "equivalent players" to count the number of significant generators. The chapter finally analyzes the impact of network constraints on the performance of deregulation. Simulations are carried on two realistic electricity transmission networks (bus systems). Our supply function equilibrium model is compared to the regulated system but also to a Cournot market. Under both networks, our supply function model performs well. It generates a social welfare that is not only better than that of a Cournot market but that is also very close to that of the regulated system. Our SFE market generates 99.9% and 99.2% of the optimal social welfare in the 30-bus and 5-bus systems respectively. With such good performance, deregulation is clearly cost effective as this small loss of social welfare is far outweighed by the cost of maintaining a centrally regulated system. In conclusion, this chapter shows that, done properly and under the appropriate market structure, deregulation benefits society without harming consumers.

# Conclusions

This thesis studies the effects of free competition in various market settings. It begins with a general quantity (Cournot) competition model with oligopolistic suppliers facing independent constraints on their production sets. The thesis then investigates a decentralized competitive mechanism to coordinate the subsidiaries of a parent company and keep their global energy consumption below a target level. The thesis finally analyzes electricity markets. It proposes a new supply function equilibrium model to represent competition and evaluates the impact of deregulation in this market. In each setting, we compare free competition and a centrally regulated mechanism in terms of market price and quantities of product sold. We also consider the performance of free competition in terms of suppliers profit, consumers surplus and social welfare. For each of these measures, we estimate the ratio of performance between free competition and the regulated setting.

The paper shows how the market structure (number of suppliers, intensity of competition, ...) affects the performance of free competition. In comparing Chapter 2 and 3, the thesis highlights how the presence of joint constraints tying suppliers together radically changes the analysis of free competition. While deregulation is never too costly when suppliers are not bound together by constraints, it can become arbitrarily costly in the presence of a joint constraint. Similarly, Chapter 4 shows that physical constraints such as those imposed by an electrical network have a significant impact on the performance of deregulation. Through rigorous analysis of complex market settings, the thesis thus contributes to deepen the understanding of free competition and its impact on market participants. It provides useful knowledge for regulators to decide whether or not a given market should be deregulated.

# Appendix A

# Proofs for Chapter 2

## A.1 Proof of Theorem 2.1

**Theorem A.1.** *In a market with differentiated substitute products, a single product per firm and separate capacity constraints for each product, colluding firms always sell less quantity of each product than if they compete freely:* $\mathbf{d}^{MP} \leq \mathbf{d}^{OP}$.

*Proof.* To prove this theorem we first formulate the oligopoly problem (OP) under capacity constraints. It can be written as:

$$\begin{aligned}
\max_{d_i} \quad & d_i \cdot \left\{ \bar{p}_i - (\ \mathbf{B}_i\ ) \cdot \begin{pmatrix} d_i \\ \mathbf{d}_{-i}^{OP} \end{pmatrix} \right\} \\
\text{s.t.} \quad & 0 \leq d_i \leq C_i \leq \bar{d}_i
\end{aligned}$$

where $\mathbf{B}_i$ denotes the row of matrix $\mathbf{B}$ corresponding to firm $i$.

Using notation $\boldsymbol{\Gamma} = \text{diag}(\mathbf{B})$, the corresponding (OP) KKT conditions are:

$$\bar{\mathbf{p}} - \mathbf{B}\mathbf{d}^{OP} - \boldsymbol{\Gamma}\mathbf{d}^{OP} - \boldsymbol{\lambda}^{OP} + \boldsymbol{\mu}^{OP} = 0 \qquad \left\{ \begin{array}{c} \lambda_i^{OP}(C_i - d_i^{OP}) = 0 \\ \lambda_i^{OP} \geq 0 \\ d_i^{OP} \leq C_i \leq \bar{d}_i \end{array} \right. \qquad \left\{ \begin{array}{c} \mu_i^{OP} d_i^{OP} = 0 \\ \mu_i^{OP} \geq 0 \\ d_i^{OP} \geq 0 \end{array} \right.$$

Similarly, we write down the monopoly problem (MP) under capacity constraints.

$$\begin{aligned} \max_{\mathbf{d}} \quad & \mathbf{d} \cdot \{\bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}\} \\ \text{s.t.} \quad & 0 \leq \mathbf{d} \leq \mathbf{C} \leq \bar{\mathbf{d}} \end{aligned}$$

The corresponding (MP) KKT conditions are:

$$\bar{\mathbf{p}} - 2\mathbf{B}\mathbf{d}^{MP} - \boldsymbol{\lambda}^{MP} + \boldsymbol{\mu}^{MP} = 0 \qquad \left\{ \begin{array}{c} (\boldsymbol{\lambda}^{MP})^T(\mathbf{C} - \mathbf{d}^{MP}) = 0 \\ \boldsymbol{\lambda}^{MP} \geq 0 \\ \mathbf{d}^{MP} \leq \mathbf{C} \leq \bar{\mathbf{d}} \end{array} \right. \qquad \left\{ \begin{array}{c} (\boldsymbol{\mu}^{MP})^T \mathbf{d}^{MP} = 0 \\ \boldsymbol{\mu}^{MP} \geq 0 \\ \mathbf{d}^{MP} \geq 0 \end{array} \right.$$

Step 1: We will prove that $\boldsymbol{\mu}^{OP} = 0$

Let us consider the problem that ignores the constraint $\mathbf{d}^{OP} \geq 0$. This suggests we ignore $\boldsymbol{\mu}^{OP}$ and the KKT conditions of problem (OP) become:

$$\bar{\mathbf{p}} - (\mathbf{B} + \boldsymbol{\Gamma})\mathbf{d}^{OP} - \boldsymbol{\lambda}^{OP} = 0 \qquad \text{or} \qquad \mathbf{d}^{OP} = (\mathbf{B} + \boldsymbol{\Gamma})^{-1}(\bar{\mathbf{p}} - \boldsymbol{\lambda}^{OP})$$

with $\mathbf{B} + \boldsymbol{\Gamma}$ being an inverse M-Matrix (see [72]).

There are two cases to distinguish.

- Either $\lambda_j^{OP} > 0$, in which case: $d_j^{OP} = C_j > 0$
- Or $\lambda_j^{OP} = 0$,

$$\begin{aligned} d_j^{OP} &= (\mathbf{B} + \boldsymbol{\Gamma})_j^{-1}(\bar{\mathbf{p}} - \boldsymbol{\lambda}^{OP}) \\ &= (- \cdots - \underbrace{+}_{jj} - \cdots -) \begin{pmatrix} \bar{p}_1 - \lambda_1^{OP} \\ \bar{p}_j \\ \bar{p}_n - \lambda_n^{OP} \end{pmatrix} \geq (- \cdots - + - \cdots -)\, \bar{\mathbf{p}} \\ d_j^{OP} &\geq (\mathbf{B} + \boldsymbol{\Gamma})_j^{-1}\, \mathbf{B}\, \bar{\mathbf{d}} \;\; = (I + \mathbf{M}\boldsymbol{\Gamma})_j^{-1}\, \bar{\mathbf{d}} \;\; > 0 \end{aligned}$$

Since $\mathbf{M}$ is an M-matrix, so is $I + \mathbf{M\Gamma}$ (see [72]). Hence $(I + \mathbf{M\Gamma})^{-1}$ has non-negative elements, and the last inequality follows from $\bar{\mathbf{d}} > 0$.

Hence, it is always the case that $\mathbf{d}^{OP} \geq 0$ even without including this constraint (i.e. the constraint that $\mathbf{d}^{OP} \geq 0$). As a result, $\boldsymbol{\mu}^{OP} = 0$.

Step 2: Similarly, we now show that $\boldsymbol{\mu}^{MP} = 0$

Following a similar thought process as before, we first consider the problem that ignores $\boldsymbol{\mu}^{MP}$ (that is, ignores the constraint $\mathbf{d}^{MP} \geq 0$). Then the KKT conditions of problem (MP) become:

$$\bar{\mathbf{p}} - 2\mathbf{B}\ \mathbf{d}^{MP} - \boldsymbol{\lambda}^{MP} = 0 \qquad \text{or} \qquad \mathbf{d}^{MP} = 1/2\ \mathbf{M}(\bar{\mathbf{p}} - \boldsymbol{\lambda}^{MP})$$

- Either $\lambda_j^{MP} > 0$, in which case: $d_j^{MP} = C_j > 0$
- Or $\lambda_j^{MP} = 0$,

$$\begin{aligned} d_j^{MP} &= 1/2\ \mathbf{M}_j\ (\bar{\mathbf{p}} - \lambda^{MP}) \\ &= (- \cdots - \underbrace{+}_{jj} - \cdots -) \begin{pmatrix} \bar{p}_1 - \lambda_1^{MP} \\ \bar{p}_j \\ \bar{p}_n - \lambda_n^{MP} \end{pmatrix} \geq 1/2\ \mathbf{M}_j\ \bar{\mathbf{p}} \\ d_j^{MP} &\geq 1/2\ \bar{d}_j \quad > 0 \end{aligned} \tag{A.1}$$

Step 3: *Characterization of* $\mathbf{d}^{OP}$

Let $K_1 = \{\text{Set of active constraints for the oligopoly problem}\} = \{i = 1, \ldots, n,\ \ \lambda_i^{OP} > 0\}$. We denote by $K_1^c$ the complement set of $K_1$ and by $\mathbf{H}_{AB}$ and $\mathbf{u}_A$ the restrictions of matrix $\mathbf{H}$ and vector $\mathbf{u}$ to rows indexed by $A$ and columns indexed by $B$. Since $K_1$ is the set of active capacity constraints for problem (OP), $\mathbf{d}^{OP} = \begin{pmatrix} d_{K_1}^{OP} \\ d_{K_1^c}^{OP} \end{pmatrix} = \begin{pmatrix} c_{K_1} \\ d_{K_1^c}^{OP} \end{pmatrix}$.

Since $\boldsymbol{\mu}^{OP} = 0$, the oligopoly KKT conditions become:

$$\bar{\mathbf{p}} - (\mathbf{B} + \boldsymbol{\Gamma})\mathbf{d}^{OP} - \boldsymbol{\lambda}^{OP} = 0$$

Restricting attention to the set $K_1^c$ of inactive constraints ($\lambda_{K_1^c}^{OP} = 0$) and noting that $\boldsymbol{\Gamma}$ disappears in off-diagonal block matrices:

$$\bar{\mathbf{p}}_{K_1^c} - \ \mathbf{B}_{K_1^c K_1}\ c_{K_1} - (\mathbf{B} + \boldsymbol{\Gamma})_{K_1^c K_1^c}\ d_{K_1^c}^{OP} = 0$$

Using the relation $\bar{\mathbf{p}}_{K_1^c} = \mathbf{B}_{K_1^c}\ \bar{\mathbf{d}}$, we get:

$$\begin{aligned} & (\mathbf{B} + \boldsymbol{\Gamma})_{K_1^c K_1^c}\ d_{K_1^c}^{OP} = \mathbf{B}_{K_1^c K_1}\ \bar{\mathbf{d}}_{K_1} + \mathbf{B}_{K_1^c K_1^c}\ \bar{\mathbf{d}}_{K_1^c} - \ \mathbf{B}_{K_1^c K_1}\ c_{K_1} \\ \Rightarrow \quad & (\mathbf{B} + \boldsymbol{\Gamma})_{K_1^c K_1^c}\ d_{K_1^c}^{OP} = \mathbf{B}_{K_1^c K_1}(\bar{\mathbf{d}}_{K_1} - c_{K_1}) + \mathbf{B}_{K_1^c K_1^c}\ \bar{\mathbf{d}}_{K_1^c} \end{aligned} \tag{A.2}$$

Clearly, on $K_1$ we have: $d_{K_1}^{OP} = c_{K_1} \geq d_{K_1}^{MP}$. Hence, to prove the Theorem above, we only need to show: $d_{K_1^c}^{OP} \geq d_{K_1^c}^{MP}$.

Step 4: *Characterization of* $\mathbf{d}^{MP}$

Let $K_2 = \{$Set of active constraints for the monopoly problem$\}$ $= \{i = 1, \ldots, n, \quad \lambda_i^{MP} > 0\}$. We denote by $K_2^c$ the complement set of $K_2$. Since $K_2$ is the set of active capacity constraints for problem (MP), $\mathbf{d}^{MP} = \begin{pmatrix} c_{K_2} \\ d_{K_2^c}^{MP} \end{pmatrix}$.

Since $\boldsymbol{\mu}^{MP} = 0$, the monopoly KKT conditions become:

$$\bar{\mathbf{p}} - 2\ \mathbf{B}\ \mathbf{d}^{MP} - \boldsymbol{\lambda}^{MP} = 0$$

Restricting attention to the set $K_2^c$ of inactive constraints ($\lambda_{K_2^c}^{MP} = 0$):

$$\bar{\mathbf{p}}_{K_2^c} - \ 2\ \mathbf{B}_{K_2^c}\ d^{MP} = 0 \tag{A.3}$$

Without loss of generality, we now assume $K_2 \subseteq K_1$ (and hence $K_2^c \supseteq K_1^c$). If there were constraints in $K_2 \setminus K_1$, we simply remove them. We show that without these constraints $d_{K_1^c}^{MP} \leq d_{K_1^c}^{OP}$ which proves that capacity constraints cannot be active on $d_{K_1^c}^{MP}$ as they are not active on $d_{K_1^c}^{OP}$.

Restricting further (A.3) to $K_1^c$ ($\subseteq K_2^c$) and splitting variables according to $K_1 \mid K_1^c$, we get:

$$\bar{\mathbf{p}}_{K_1^c} - \ 2\ \mathbf{B}_{K_1^c K_1} \begin{pmatrix} c_{K_2} \\ d_{K_1 \setminus K_2}^{MP} \end{pmatrix} - 2\ \mathbf{B}_{K_1^c K_1^c}\ d_{K_1^c}^{MP} = 0$$

Using the relation $\bar{\mathbf{p}}_{K_1^c} = \mathbf{B}_{K_1^c}\ \bar{\mathbf{d}}$, we get:

$$2\ \mathbf{B}_{K_1^c K_1^c}\ d_{K_1^c}^{MP} = \mathbf{B}_{K_1^c K_1}\ \bar{\mathbf{d}}_{K_1} + \mathbf{B}_{K_1^c K_1^c}\ \bar{\mathbf{d}}_{K_1^c} - \ 2\ \mathbf{B}_{K_1^c K_1} \begin{pmatrix} c_{K_2} \\ d_{K_1 \setminus K_2}^{MP} \end{pmatrix}$$

$$\Rightarrow \quad 2\ \mathbf{B}_{K_1^c K_1^c}\ d_{K_1^c}^{MP} = \mathbf{B}_{K_1^c K_1} \left( \bar{\mathbf{d}}_{K_1} - \begin{matrix} 2\ c_{K_2} \\ 2\ d_{K_1 \setminus K_2}^{MP} \end{matrix} \right) + \mathbf{B}_{K_1^c K_1^c}\ \bar{\mathbf{d}}_{K_1^c} \tag{A.4}$$

<u>Step 5:</u> $\mathbf{d}^{OP} \geq \mathbf{d}^{MP}$

As shown in (A.1), for all $j \in K_2^c$, $d_j^{MP} \geq 1/2\ \bar{d}_j$. In particurlar:

$$2\ d_{K_1 \setminus K_2}^{MP} \ \geq \ \bar{\mathbf{d}}_{K_1 \setminus K_2} \ \geq \ c_{K_1 \setminus K_2} \tag{A.5}$$

$$2\ d_{K_1^c}^{MP} \ \geq \ \bar{\mathbf{d}}_{K_1^c} \tag{A.6}$$

On the other hand, combining (A.2) and (A.4), we have:

$$
\begin{aligned}
(\mathbf{B}+\mathbf{\Gamma})_{K_1^c K_1^c}\, d^{OP}_{K_1^c} - \mathbf{B}_{K_1^c K_1}(\bar{\mathbf{d}}_{K_1} - c_{K_1}) \;&=\; 2\,\mathbf{B}_{K_1^c K_1^c}\, d^{MP}_{K_1^c} - \mathbf{B}_{K_1^c K_1}\begin{pmatrix} \bar{\mathbf{d}}_{K_1} - & 2\,c_{K_2} \\ & 2\,d^{MP}_{K_1\setminus K_2} \end{pmatrix} \\
\Rightarrow (\mathbf{B}+\mathbf{\Gamma})_{K_1^c K_1^c}\, d^{OP}_{K_1^c} \;&=\; 2\,\mathbf{B}_{K_1^c K_1^c}\, d^{MP}_{K_1^c} + \mathbf{B}_{K_1^c K_1}\underbrace{\begin{pmatrix} 2\,c_{K_2} & - & c_{K_2} \\ 2\,d^{MP}_{K_1\setminus K_2} & & c_{K_1\setminus K_2} \end{pmatrix}}_{\geq 0 \quad \text{using } (A.5)} \\
\Rightarrow (\mathbf{B}+\mathbf{\Gamma})_{K_1^c K_1^c}\, d^{OP}_{K_1^c} \;&\geq\; 2\,\mathbf{B}_{K_1^c K_1^c}\, d^{MP}_{K_1^c}
\end{aligned} \tag{A.7}
$$

Finally, let's assume there exist $i \in K_1^c$ such that $d_i^{OP} < d_i^{MP}$. Denoting $\{s_1, \cdots, s_f\}$ the indices of $K_1^c$, let's expand the i-th row of (A.7):

$$
(b_{is_1} \cdots 0 \cdots b_{is_f})\ \underbrace{d^{OP}_{K_1^c}}_{\leq\, \bar{\mathbf{d}}_{K_1^c}} + 2\,b_{ii}\ \underbrace{d_i^{OP}}_{<\, d_i^{MP}} \;\geq\; (\,b_{is_1} \cdots 0 \cdots b_{is_f})\ \underbrace{2\,d^{MP}_{K_1^c}}_{\substack{\geq\, \bar{\mathbf{d}}_{K_1^c} \\ \text{using } (A.6)}} + 2\,b_{ii}\,d_i^{MP}
$$

Since all the coefficients $b_{ij}$ are non-negative, this is a contradiction.
We just showed that $d^{MP}_{K_1^c} \leq d^{OP}_{K_1^c}$, leading to $d^{MP} \leq d^{OP}$.

□

## A.2 Proof of Step 1 for Theorem 2.3

Ignoring $\boldsymbol{\mu}^{SP}$, the KKT conditions of problem (SP) become:

$$
\bar{\mathbf{p}} - \mathbf{B}\ \mathbf{d}^{SP} - \boldsymbol{\lambda}^{SP} = 0 \qquad \text{or} \qquad \mathbf{d}^{SP} = \mathbf{M}\ (\bar{\mathbf{p}} - \boldsymbol{\lambda}^{SP})
$$

- Either $\lambda_j^{SP} > 0$, in which case: $d_j^{SP} = C_j > 0$

- Or $\lambda_j^{SP} = 0$,

$$\begin{aligned}
d_j^{SP} &= \mathbf{M}_j\ (\bar{\mathbf{p}} - \boldsymbol{\lambda}^{SP}) \\
&= (- \cdots - \underbrace{+}_{jj} - \cdots -) \begin{pmatrix} \bar{p}_1 - \lambda_1^{SP} \\ \bar{p}_j \\ \bar{p}_n - \lambda_n^{SP} \end{pmatrix} \geq \mathbf{M}_j\ \bar{\mathbf{p}} \\
d_j^{SP} &\geq \bar{d}_j \quad > 0
\end{aligned}$$

## A.3 Calculations for Theorem 2.4

In the uniform case, matrix $\mathbf{M}$ can be written as:

$$\begin{aligned}
\mathbf{M} = \begin{pmatrix} 1 & -\alpha & \dots & -\alpha \\ -\alpha & \ddots & & \vdots \\ \vdots & & \ddots & -\alpha \\ -\alpha & \dots & -\alpha & 1 \end{pmatrix} &= (1+\alpha)I - \alpha H \\
&= \Delta \begin{pmatrix} 1+\alpha-n\alpha & 0 & \dots & 0 \\ 0 & 1+\alpha & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1+\alpha \end{pmatrix} \Delta^T
\end{aligned}$$

Inverting $\mathbf{M}$, we get matrix $\mathbf{B}$:

$$\begin{aligned}
\mathbf{B} &= \frac{1}{1+\alpha}(I - \frac{\alpha}{1+\alpha}H)^{-1} \\
&= \frac{1}{1+\alpha}\left[I + \frac{\alpha}{1+\alpha}(1 + \frac{\alpha}{1+\alpha}n + \cdots)H\right] \\
&= \frac{1}{1+\alpha}\left[I + \frac{\alpha}{1+\alpha-n\alpha}H\right]
\end{aligned}$$

This allows us to compute:

$$\mathbf{\Gamma} = \text{diag}(\mathbf{B}) = \frac{1 + 2\ \alpha - n\alpha}{(1+\alpha)(1+\alpha-n\alpha)} I$$

On the other hand, diagonalizing $\mathbf{B}$ as we did with $\mathbf{M}$:

$$\mathbf{B} = \Delta \begin{pmatrix} \frac{1}{1+\alpha-n\alpha} & 0 & \dots & 0 \\ 0 & \frac{1}{1+\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{1+\alpha} \end{pmatrix} \Delta^T$$

We are now able to compute the diverse component of the surplus ratio.

$$I + \mathbf{M\Gamma} = \Delta \begin{pmatrix} \frac{2+3\alpha-n\alpha}{1+\alpha} & 0 & \dots & 0 \\ 0 & \frac{2+3\alpha-2n\alpha}{1+\alpha-n\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{2+3\alpha-2n\alpha}{1+\alpha-n\alpha} \end{pmatrix} \Delta^T$$

$$(I + \mathbf{M\Gamma})^{-1} = \Delta \begin{pmatrix} \frac{1+\alpha}{2+3\alpha-n\alpha} & 0 & \dots & 0 \\ 0 & \frac{1+\alpha-n\alpha}{2+3\alpha-2n\alpha} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \frac{1+\alpha-n\alpha}{2+3\alpha-2n\alpha} \end{pmatrix} \Delta^T$$

Let's call $\breve{\mathbf{d}}$ the vector whose components are the eigenvectors of $\mathbf{M}$, and $[\breve{\rho}_1, \breve{\rho}_2]$ the two eigenvalues of: $(I + \mathbf{\Gamma M})^{-1}\ \mathbf{\Gamma}\ (I + \mathbf{M\Gamma})^{-1}$.

- $\breve{\rho}_1 = \frac{(1+\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-n\alpha)^2(1+\alpha-n\alpha)}$
- $\breve{\rho}_2 = \frac{(1+\alpha-n\alpha)(1+2\alpha-n\alpha)}{(2+3\alpha-2n\alpha)^2(1+\alpha)}$

The ratio of profits becomes:

$$\frac{\Pi(OP)}{\Pi(MP)} = \frac{4\ (\breve{\rho}_1 \breve{d}_1^2 + \breve{\rho}_2 \sum_{i=2}^n \breve{d}_i^2)}{\frac{1}{1+\alpha-n\alpha}\breve{d}_1^2 + \frac{1}{1+\alpha}\sum_{i=2}^n \breve{d}_i^2}$$

## A.4 Proof of Lemma B.1

**Lemma A.1.** *For a symmetric inverse M-matrix* $\mathbf{B}$ *and a vector* $\mathbf{d}$ *with all component positive, the following inequality holds:*

$$\|\mathbf{d}\|_{\mathbf{B}}^2 \leq (1 + r \cdot (nm - 1)) \; \|\mathbf{d}\|_{\mathbf{B}\mathbf{Bdiag}}^2$$

*where* $r$ *is the market power.*

*Proof.* Since $\mathbf{B}$ is an inverse M-matrix, Ostrowski shows in [94] that:

$$B_{ij}^{kl} \leq r_{kl} B_{ij}^{ij} \qquad \textit{and} \qquad B_{ij}^{kl} = B_{kl}^{ij} \leq r_{ij} B_{kl}^{kl}$$

Introducing $r = \max_{kl} r_{kl}$, we have: $B_{ij}^{kl} \leq r\sqrt{B_{ij}^{ij} B_{kl}^{kl}}$.

Hence, we can write:

$$\begin{aligned}
\|\mathbf{d}\|_{\mathbf{B}}^2 &\leq \mathbf{d}^T \begin{pmatrix} B_{11}^{11} & \dots & r\sqrt{B_{ij}^{ij} B_{kl}^{kl}} \\ \vdots & \ddots & \vdots \\ r\sqrt{B_{ij}^{ij} B_{kl}^{kl}} & \dots & B_{nm}^{nm} \end{pmatrix} \mathbf{d} \\
&= \mathbf{d}^T \begin{pmatrix} rB_{11}^{11} & \dots & r\sqrt{B_{ij}^{ij} B_{kl}^{kl}} \\ \vdots & \ddots & \vdots \\ r\sqrt{B_{ij}^{ij} B_{kl}^{kl}} & \dots & rB_{nm}^{nm} \end{pmatrix} \mathbf{d} \\
&\quad + \mathbf{d}^T \begin{pmatrix} (1-r)B_{11}^{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (1-r)B_{nm}^{nm} \end{pmatrix} \mathbf{d}
\end{aligned}$$

We denote the diagonal matrix corresponding to the diagonal of matrix $\mathbf{B}$ by:

$$\mathbf{\Gamma} = \mathrm{diag}(B_{11}^{11}, \cdots, B_{nm}^{nm})$$

We obtain:

$$\|\mathbf{d}\|_{\mathbf{B}}^2 \leq r\ \mathbf{d}^T\sqrt{\mathbf{\Gamma}} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \sqrt{\mathbf{\Gamma}}\ \mathbf{d} + (1-r)\ \mathbf{d}^T\mathbf{\Gamma}\mathbf{d}$$

Since $\mathbf{H} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$ has two eigenvalues 0 and $nm$, we have $\mathbf{d}^T\mathbf{H}\mathbf{d} \leq nm\ \|\mathbf{d}\|^2$ for all $\mathbf{d}$.

$$\begin{aligned} \|\mathbf{d}\|_{\mathbf{B}}^2 &\leq r \cdot nm\ \mathbf{d}^T\mathbf{\Gamma}\mathbf{d} + (1-r)\ \mathbf{d}^T\mathbf{\Gamma}\mathbf{d} \\ &\leq (1 + r\cdot(nm-1))\ \|\mathbf{d}\|_{\mathbf{BBdiag}}^2 \end{aligned}$$

□

## A.5 Derivation of oligopoly variational inequality

At a Nash equilibrium solution, the optimization problem facing a single firm is:

$$\max_{\mathbf{d}_i} \quad \mathbf{d}_i \cdot \left\{ \bar{\mathbf{p}}_i - \begin{pmatrix} B_{i1} \\ \vdots \\ B_{im} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{d}_i \\ \mathbf{d}_{-i}^{OP} \end{pmatrix} \right\} \tag{A.8}$$

$$\text{s.t.} \qquad \mathbf{d}_i \in K_i$$

This problem is a maximization of a concave objective function over a convex set, it is a convex problem. A general convex problem of the form:

$$\begin{aligned} \max_x \quad & F(x) \\ \text{s.t.} \quad & x \in K \end{aligned}$$

with a concave objective $F(x)$ is equivalent (see [85], [112]) to the variational inequality problem:

$$\text{Find } x_0 \in K : \quad -\nabla F(x_0)\ (x - x_0) \geq 0 \qquad \forall x \in K$$

Applying this to (A.8), we obtain for each firm $i$:

$$\text{Find } \mathbf{d}_i^{OP} \in K_i : \quad \left\{-\bar{\mathbf{p}}_i + \mathbf{B}_i \cdot \mathbf{d}^{OP} + \mathbf{B}_i^i \cdot \mathbf{d}_i^{OP}\right\}^T (\mathbf{d}_i - \mathbf{d}_i^{OP}) \geq 0 \qquad \forall \mathbf{d}_i \in K_i$$

where $\mathbf{B}_i$ denotes the rows of matrix $\mathbf{B}$ corresponding to firm $i$.

Now, since the constraint set of each firm $i$ is independent of the quantities chosen by other firms, it is equivalent to satisfy every one of these variational inequalitites (for firm $i$) or to satisfy the sum of these inequalities. Clearly, if $\mathbf{d}^{OP}$ satisfies all these inequalities it satisfies the sum of the inequalities. On the other hand if $\mathbf{d}^{OP}$ satisfies the sum of the inequalities, by choosing $\mathbf{d} = (\mathbf{d}_i, \mathbf{d}_{-i}^{OP})$ for all $\mathbf{d}_i \in K_i$, it is easy to check that it will satisfy every variational inequality separately as well. The sum of these inequalities is exactly the variational inequality used in this chapter:

$$\text{Find } \mathbf{d}^{OP} \in K : \quad \left\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \mathbf{B}^{\mathbf{Bdiag}} \cdot \mathbf{d}^{OP}\right\}^T (\mathbf{d} - \mathbf{d}^{OP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

# Appendix B

# Proofs for Chapter 3

## B.1 Calculations and proofs for Section 3.3

### B.1.1 KKT conditions

The KKT conditions of the different problems are summarized as follows:

$$\text{SMAX:} \qquad \begin{cases} \bar{\mathbf{p}} - B\mathbf{d}^{SMAX} - \mu\mathbf{e} + \boldsymbol{\lambda} = \mathbf{0} & \\ \mathbf{e}^T\mathbf{d}^{SMAX} \leq c & \\ \mu(c - \mathbf{e}^T\mathbf{d}) = 0 & \\ \lambda_i d_i^{SMAX} = 0 & \forall i = 1, 2, ..., n \\ \lambda_i \geq 0, \mu \geq 0, d_i^{SMAX} \geq 0 & \end{cases}$$

$$\text{CP:} \qquad \begin{cases} \bar{\mathbf{p}} - 2B\mathbf{d}^{CP} - \mu\mathbf{e} + \boldsymbol{\lambda} = \mathbf{0} & \\ \mathbf{e}^T\mathbf{d}^{CP} \leq c & \\ \mu(c - \mathbf{e}^T\mathbf{d}) = 0 & \\ \lambda_i d_i^{CP} = 0 & \forall i = 1, 2, ..., n \\ \lambda_i \geq 0, \mu \geq 0, d_i^{CP} \geq 0 & \end{cases}$$

$$
\text{UNE:} \qquad \begin{cases} \bar{\mathbf{p}} - (\mathbf{B}+\Gamma)\mathbf{d}^{UNE} - \mu\mathbf{e} + \boldsymbol{\lambda} = \mathbf{0} \\ \mathbf{e}^T\mathbf{d}^{UNE} \leq c \\ \mu(c - \mathbf{e}^T\mathbf{d}) = 0 \\ \lambda_i \; d_i^{UNE} = 0 & \forall i = 1,2,...,n \\ \lambda_i \geq 0, \mu \geq 0, d_i^{UNE} \geq 0 \end{cases}
$$

The set of oligopoly equilibria satisfies the following Cauchy variational inequality:

$$
\left(-\bar{p}_i + (\mathbf{B}+\Gamma)_i\mathbf{d}^{OP}\right)\left(d_i - d_i^{OP}\right) \geq 0 \qquad \forall d_i \in K(\mathbf{d}_{-i}^{OP})
$$

where $K(\mathbf{d}_{-i}^{OP}) = \left\{d_i | d_i + \sum_{j\neq i}^{n} d_j^{OP} \leq C, \; d_i \geq 0\right\}$.

### B.1.2 Proofs

**Theorem B.1.** *In the absence of constraints, the loss of profit resulting from free competition between subsidiaries of a company is bounded by:*

$$
\Pi(OP)/\Pi(CP) \geq \max\left\{\frac{2}{2 + r\cdot(n-1)}; \frac{3}{4 + r\cdot(n-1)}\right\}
$$

*where $r$ is the diversion ratio, $n$ is the number of subsidiaries.*

The bound is composed of two parts. The first part dominates when $r \cdot (n-1) \leq 2$, the second part dominates otherwise.

In order to prove the theorem, we need to use an intermediate inequality. To shorten notations, we denote by $||\mathbf{d}||_{\mathbf{B}}^2 = \mathbf{d}^T\,\mathbf{B}\mathbf{d}$, and $||\mathbf{d}||_{\Gamma}^2 = \mathbf{d}^T\,\Gamma\mathbf{d}$.

**Lemma B.1.** *For a symmetric inverse M-matrix* $\mathbf{B}$ *and a vector* $\mathbf{d}$ *with all component positive, the following inequality holds:*

$$
||\mathbf{d}||_{\mathbf{B}}^2 \leq (1 + r\cdot(n-1))\;||\mathbf{d}||_{\Gamma}^2
$$

*where $r$ is the diversion ratio.*

*Proof.* Since $\mathbf{B}$ is an inverse M-matrix, Ostrowski shows in [94] that:

$$B_{ij} \leq r_j \ B_{ii} \qquad and \qquad B_{ij} = B_{ji} \leq r_i \ B_{jj}$$

Introducing $r = \max_i r_i$, we have: $B_{ij} \leq r\sqrt{B_{ii}B_{jj}}$.

Hence, we can write:

$$\begin{aligned}
\|\mathbf{d}\|^2_{\mathbf{B}} &\leq \mathbf{d}^T \begin{pmatrix} B_{11} & \dots & r\sqrt{B_{ii}B_{jj}} \\ \vdots & \ddots & \vdots \\ r\sqrt{B_{ii}B_{jj}} & \dots & B_{nn} \end{pmatrix} \mathbf{d} \\
&= \mathbf{d}^T \begin{pmatrix} rB_{11} & \dots & r\sqrt{B_{ii}B_{jj}} \\ \vdots & \ddots & \vdots \\ r\sqrt{B_{ii}B_{jj}} & \dots & rB_{nn} \end{pmatrix} \mathbf{d} + \mathbf{d}^T \begin{pmatrix} (1-r)B_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (1-r)B_{nn} \end{pmatrix} \mathbf{d}
\end{aligned}$$

We denote the diagonal matrix corresponding to the diagonal of matrix $\mathbf{B}$ by:

$$\Gamma = \mathrm{diag}(B_{11}, \cdots, B_{nn})$$

We obtain:

$$\|\mathbf{d}\|^2_{\mathbf{B}} \leq r \ \mathbf{d}^T\sqrt{\Gamma} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix} \sqrt{\Gamma} \ \mathbf{d} + (1-r) \ \mathbf{d}^T\Gamma\mathbf{d}$$

Since $\mathbf{H} = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$ has two eigenvalues 0 and $n$, we have $\mathbf{d}^T\mathbf{H}\mathbf{d} \leq n \ \|\mathbf{d}\|^2$ for all $\mathbf{d}$.

$$\begin{aligned} \|\mathbf{d}\|_{\mathbf{B}}^2 &\leq r \cdot n \ \mathbf{d}^T \Gamma \mathbf{d} + (1-r) \ \mathbf{d}^T \Gamma \mathbf{d} \\ &\leq (1 + r \cdot (n-1)) \ \|\mathbf{d}\|_{\Gamma}^2 \end{aligned}$$

□

We are now ready to prove Theorem B.1:

*Proof.* The variational inequality satisfied at the oligopoly solution is:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \Gamma \cdot \mathbf{d}^{OP}\}^T (\mathbf{d} - \mathbf{d}^{OP}) \geq 0 \qquad \forall \mathbf{d} \in K \tag{B.1}$$

Since by definition the centrally coordinated solution must be feasible as well (i.e. $\mathbf{d}^{CP} \in K$), the first order optimality conditions evaluated at the centrally coordinated solution are:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{OP} + \Gamma \cdot \mathbf{d}^{OP}\}^T (\mathbf{d}^{CP} - \mathbf{d}^{OP}) \geq 0$$

Denoting by $\Pi = \mathbf{d} \cdot \{\bar{\mathbf{p}} - \mathbf{B} \cdot \mathbf{d}\}$ the profit, we get:

$$\Pi(OP) - \|\mathbf{d}^{OP}\|_{\Gamma}^2 - \bar{\mathbf{p}}^T \mathbf{d}^{CP} + (\mathbf{d}^{OP})^T \mathbf{B} \mathbf{d}^{CP} + (\mathbf{d}^{OP})^T \Gamma \mathbf{d}^{CP} \geq 0$$

By adding and substracting $(\mathbf{d}^{CP})^T B \mathbf{d}^{CP}$, we obtain:

$$\Pi(OP) - \Pi(CP) + (\mathbf{d}^{OP})^T \mathbf{B} \mathbf{d}^{CP} - \|\mathbf{d}^{CP}\|_{\mathbf{B}}^2 - \|\mathbf{d}^{OP}\|_{\Gamma}^2 + (\mathbf{d}^{OP})^T \Gamma \mathbf{d}^{CP} \geq 0 \tag{B.2}$$

<u>Part 1:</u> Let's first prove the second part of the bound, namely $\Pi(OP)/\Pi(CP) \geq \frac{3}{4+r\cdot(n-1)}$.

Since both $\mathbf{B}$ and $\Gamma$ are positive definite matrices we can use the following bounds:

- $(\mathbf{d}^{OP})^T \mathbf{B} \mathbf{d}^{CP} \leq 1/3 \|\mathbf{d}^{OP}\|_{\mathbf{B}}^2 + 3/4 \|\mathbf{d}^{CP}\|_{\mathbf{B}}^2$

- $(\mathbf{d}^{OP})^T\Gamma\mathbf{d}^{CP} \leq \|\mathbf{d}^{OP}\|_\Gamma^2 + 1/4\|\mathbf{d}^{CP}\|_\Gamma^2$

Introducing these bounds into the variational inequality:

$$\Pi(OP) - \Pi(CP) + 1/3\|\mathbf{d}^{OP}\|_\mathbf{B}^2 \underbrace{-1/4\|\mathbf{d}^{CP}\|_\mathbf{B}^2 + 1/4\|\mathbf{d}^{CP}\|_\Gamma^2}_{=-1/4\|\mathbf{d}^{CP}\|_{\mathbf{B}-\Gamma}^2 \leq 0} \geq 0$$

Since we assumed $0 \in K$, we can plug the feasible point 0 into the variational inequality (B.1) to get:

$$\Pi(OP) \geq \|\mathbf{d}^{OP}\|_\Gamma^2 \tag{B.3}$$

Using Lemma B.1 to upper bound $\|\mathbf{d}^{OP}\|_\mathbf{B}^2$, we finally get:

$$\begin{aligned} \Pi(OP) - \Pi(CP) + 1/3(1 + r\cdot(n-1))\ \|\mathbf{d}^{OP}\|_\Gamma^2 &\geq 0 \\ \Pi(OP) - \Pi(CP) + 1/3(1 + r\cdot(n-1))\ \Pi(OP) &\geq 0 \end{aligned}$$

This inequality is equivalent to:

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{3}{4 + r\cdot(n-1)}$$

<u>Part 2:</u> Let's now prove the first part of the bound $\Pi(OP)/\Pi(CP) \geq \frac{2}{2+r\cdot(n-1)}$.

Using symmetry of the matrix $\mathbf{B}$, we can rewrite equation (B.2) as follow:

$$\Pi(OP) - \Pi(CP) + (\mathbf{d}^{OP})^T\Gamma(\mathbf{d}^{CP} - \mathbf{d}^{OP}) - (\mathbf{d}^{CP})^T\mathbf{B}(\mathbf{d}^{CP} - \mathbf{d}^{OP}) \geq 0$$

We can decompose this expression in two different ways:

- $$\Pi(OP) - \Pi(CP) + \underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{CP})^T \Gamma (\mathbf{d}^{CP} - \mathbf{d}^{OP})}_{\leq 0} - (\mathbf{d}^{CP})^T (\mathbf{B} - \Gamma)(\mathbf{d}^{CP} - \mathbf{d}^{OP}) \geq 0 \tag{B.4}$$

- $$\Pi(OP) - \Pi(CP) + \underbrace{(\mathbf{d}^{OP} - \mathbf{d}^{CP})^T \mathbf{B} (\mathbf{d}^{CP} - \mathbf{d}^{OP})}_{\leq 0} - (\mathbf{d}^{OP})^T (\mathbf{B} - \Gamma)(\mathbf{d}^{CP} - \mathbf{d}^{OP}) \geq 0 \tag{B.5}$$

Combining these two expressions (1/2 Inequality(B.4) + 1/2 Inequality(B.5)) and leaving out the non-positive terms, we get:

$$\Pi(OP) - \Pi(CP) - 1/2 (\mathbf{d}^{CP} + \mathbf{d}^{OP})^T (\mathbf{B} - \Gamma)(\mathbf{d}^{CP} - \mathbf{d}^{OP}) \geq 0$$

$$\Pi(OP) - \Pi(CP) \underbrace{- 1/2\ (\mathbf{d}^{CP})^T (\mathbf{B} - \Gamma)\mathbf{d}^{CP}}_{\leq 0} + 1/2\ (\mathbf{d}^{OP})^T (\mathbf{B} - \Gamma)\mathbf{d}^{OP} \geq 0$$

Using Lemma B.1, we can write:

$$\|\mathbf{d}^{OP}\|^2_{(\mathbf{B}-\Gamma)} = \|\mathbf{d}^{OP}\|^2_{\mathbf{B}} - \|\mathbf{d}^{OP}\|^2_{\Gamma} \leq r \cdot (n-1) \|\mathbf{d}^{OP}\|^2_{\Gamma}$$

With inequality (B.3), we finally obtain:

$$\Pi(OP) - \Pi(CP) + 1/2\ r \cdot (n-1) \Pi(OP) \geq 0$$

This leads to:

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{2}{2 + r \cdot (n-1)}$$

□

**Theorem B.2.** *In the absence of constraint, the loss of social surplus resulting from free competition between subsidiaries of a company is at most 1/4:* $TS(OP) \geq TS(CP) = \frac{3}{4} TS(SMAX)$. *Equality is achieved when subsidiaries are independent.*

*Proof.*

$$
\begin{aligned}
&TS(OP) - TS(CP)\\
&= (\mathbf{d}^{OP})^T \left( B\bar{\mathbf{d}} - \frac{1}{2}B\mathbf{d}^{OP} \right) - (\mathbf{d}^{CP})^T \left( B\bar{\mathbf{d}} - \frac{1}{2}B\mathbf{d}^{CP} \right)\\
&= (\bar{\mathbf{d}} - \frac{\mathbf{d}^{OP} + \mathbf{d}^{CP}}{2})^T B(\mathbf{d}^{OP} - \mathbf{d}^{CP})\\
&\text{(Since } \mathbf{d}^{CP} = \frac{1}{2}\bar{\mathbf{d}}, \mathbf{d}^{OP} = (\mathbf{B} + \Gamma)^{-1}B\bar{\mathbf{d}})\\
&= \left( \bar{\mathbf{d}} - \frac{(\mathbf{B} + \Gamma)^{-1}B\bar{\mathbf{d}} + \frac{1}{2}\bar{\mathbf{d}}}{2} \right)^T B \left( (\mathbf{B} + \Gamma)^{-1}B\bar{\mathbf{d}} - \frac{1}{2}\bar{\mathbf{d}} \right)\\
&= \bar{\mathbf{d}}^T \left( \frac{3}{4}\mathbf{I} - \frac{1}{2}B(\mathbf{B} + \Gamma)^{-1} \right) B \left( (\mathbf{B} + \Gamma)^{-1}B - \frac{1}{2}\mathbf{I} \right) \bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T \left( \frac{3}{2}B - B(\mathbf{B} + \Gamma)^{-1}B \right) \left( (\mathbf{B} + \Gamma)^{-1}B - \frac{1}{2}\mathbf{I} \right) \bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T \left( \frac{3}{2}B(\mathbf{B} + \Gamma)^{-1}B - B(\mathbf{B} + \Gamma)^{-1}B(\mathbf{B} + \Gamma)^{-1}B \right.\\
&\left. -\frac{3}{4}\mathbf{B} + \frac{1}{2}B(\mathbf{B} + \Gamma)^{-1}B \right) \bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T \left( 2B(\mathbf{B} + \Gamma)^{-1}B - B(\mathbf{B} + \Gamma)^{-1}B(\mathbf{B} + \Gamma)^{-1}B - \frac{3}{4}B \right) \bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T (B(\mathbf{B} + \Gamma)^{-1}\mathbf{B} + \mathbf{B}(\mathbf{B} + \Gamma)^{-1}\Gamma(\mathbf{B} + \Gamma)^{-1}B - \frac{3}{4}B)\bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T (B - B(\mathbf{B} + \Gamma)^{-1}\Gamma + B(\mathbf{B} + \Gamma)^{-1}\Gamma(\mathbf{B} + \Gamma)^{-1}B - \frac{3}{4}B)\bar{\mathbf{d}}\\
&= \frac{1}{2}\bar{\mathbf{d}}^T (\frac{1}{4}B - B(\mathbf{B} + \Gamma)^{-1}\Gamma + B(\mathbf{B} + \Gamma)^{-1}\Gamma(\mathbf{B} + \Gamma)^{-1}B)\bar{\mathbf{d}}\\
&\text{(since } \bar{\mathbf{d}}^T B\bar{\mathbf{d}} \geq \bar{\mathbf{d}}^T \Gamma\bar{\mathbf{d}})\\
&\geq \frac{1}{2}\bar{\mathbf{d}}^T (\frac{1}{4}\Gamma - B(\mathbf{B} + \Gamma)^{-1}\Gamma + B(\mathbf{B} + \Gamma)^{-1}\Gamma(\mathbf{B} + \Gamma)^{-1}B)\bar{\mathbf{d}}
\end{aligned}
$$

let $\bar{\mathbf{d}}^T\Gamma^{\frac{1}{2}} = x^T$, and $\bar{\mathbf{d}}^T B(\mathbf{B}+\Gamma)^{-1}\Gamma^{\frac{1}{2}} = y^T$. We have

$$\begin{aligned}
TS(OP) - TS(CP) &\geq \frac{1}{2}\bar{\mathbf{d}}^T(\frac{1}{4}\Gamma - B(\mathbf{B}+\Gamma)^{-1}\Gamma + B(\mathbf{B}+\Gamma)^{-1}\Gamma(\mathbf{B}+\Gamma)^{-1}B)\bar{\mathbf{d}} \\
&= \frac{1}{2}(\frac{1}{4}x^T x - x^T y + y^T y) \\
&= \frac{1}{2}(\frac{1}{2}x - y)^T(\frac{1}{2}x - y) \\
&\geq 0
\end{aligned}$$

Equality holds when $B = \Gamma$ and $\frac{1}{2}x = y$. □

### B.1.3 Calculations for the duopoly case

$\mathbf{B} = \begin{bmatrix} \beta_1 & \alpha \\ \alpha & \beta_2 \end{bmatrix}$, $\beta_1 \geq \beta_2 > \alpha > 0$

**Centrally coordinated problem**

- case 1: (constraint is not tight) $\mathbf{d}^{SO} = \frac{1}{2}\bar{\mathbf{d}} = \begin{bmatrix} \frac{\beta_2\bar{p}_1 - \alpha\bar{p}_2}{2(\beta_1\beta_2-\alpha^2)} \\ \frac{\beta_1\bar{p}_2 - \alpha\bar{p}_1}{2(\beta_1\beta_2-\alpha^2)} \end{bmatrix} > \mathbf{0}$ (by assumption)

  if $\mathbf{e}^T\bar{\mathbf{d}}^{SO} = \frac{(\beta_2-\alpha)\bar{p}_1+(\beta_1-\alpha)\bar{p}_2}{2(\beta_1\beta_2-\alpha^2)} \leq c$

- case 2: (constraint is tight) $\mathbf{d}^{SO} = \begin{bmatrix} \frac{2c(\beta_2-\alpha)-\bar{p}_2+\bar{p}_1}{2(\beta_1+\beta_2-2\alpha)} \\ \frac{2c(\beta_1-\alpha)-\bar{p}_1+\bar{p}_2}{2(\beta_1+\beta_2-2\alpha)} \end{bmatrix}$

  if

  $2c(\beta_1-\alpha) - \bar{p}_1 + \bar{p}_2 \geq 0$ and $2c(\beta_2-\alpha) - \bar{p}_2 + \bar{p}_1 \geq 0$

  and

  $\lambda = \bar{p}_1 - 2\alpha c - 2(\beta_1-\alpha)d_1 = \frac{(\beta_2-\alpha)\bar{p}_1+(\beta_1-\alpha)\bar{p}_2-2c(\beta_1\beta_2-\alpha^2)}{\beta_1+\beta_2-2\alpha} \geq 0$

- case 3: $\mathbf{d}^{SO} = \begin{bmatrix} c \\ 0 \end{bmatrix}$

  if $\lambda = \bar{p}_1 - 2\beta_1 c \geq 0$ and $\mu_2 = \bar{p}_1 - \bar{p}_2 + 2c(\alpha - \beta_1) \geq 0$

- case 4: $\mathbf{d}^{SO} = \begin{bmatrix} 0 \\ c \end{bmatrix}$

  if $\lambda = \bar{p}_2 - 2\beta_2 c \geq 0$ and $\mu_1 = \bar{p}_2 - \bar{p}_1 + 2c(\alpha - \beta_2) \geq 0$

$\mathbf{d}^{SO}$ is unique as the above 4 cases are mutually exclusive.

### Oligopoly problem

- case 1:(constraint is not tight)

  $\mathbf{d}^{NE} = \begin{bmatrix} \frac{2\beta_2\bar{p}_1 - \alpha\bar{p}_2}{4\beta_1\beta_2 - \alpha^2} \\ \frac{2\beta_1\bar{p}_2 - \alpha\bar{p}_1}{4\beta_1\beta_2 - \alpha^2} \end{bmatrix}$

  if $\frac{(2\beta_2-\alpha)\bar{p}_1 + (2\beta_1-\alpha)\bar{p}_2}{4\beta_1\beta_2-\alpha^2} \leq c$

- case 2:(constraint is tight) $\mathbf{d}^{NE}$ satisfies the following condition:

  1. $d_1 + d_2 = c$;
  2. $-\bar{p}_1 + 2\beta_1 d_1 + \alpha d_2 \leq 0$; if $d_1 > 0$

     $-\bar{p}_2 + 2\beta_2 d_2 + \alpha d_1 \leq 0$; if $d_2 > 0$
  3. $d_1 \geq 0$ and $d_2 \geq 0$

### Uniform Nash equilibrium

- case 1: $\mathbf{d}^{UNE} = \mathbf{d}^{NE}$when constraint is not tight;

- case 2: (constraint is tight) $\mathbf{d}^{UNE} = \begin{bmatrix} \frac{c(2\beta_2-\alpha)-\bar{p}_2+\bar{p}_1}{2(\beta_1+\beta_2-\alpha)} \\ \frac{c(2\beta_1-\alpha)-\bar{p}_1+\bar{p}_2}{2(\beta_1+\beta_2-\alpha)} \end{bmatrix}$

  if

  $c(2\beta_1 - \alpha) - \bar{p}_1 + \bar{p}_2 \geq 0$ and $c(2\beta_2 - \alpha) - \bar{p}_2 + \bar{p}_1 \geq 0$

  and

  $\lambda = \bar{p}_1 - \alpha c - (2\beta_1 - \alpha)d_1 = \frac{(2\beta_2-\alpha)\bar{p}_1+(2\beta_1-\alpha)\bar{p}_2-c(4\beta_1\beta_2-\alpha^2)}{2(\beta_1+\beta_2-\alpha)} \geq 0$

- case 3: $\mathbf{d}^{UNE} = \begin{bmatrix} c \\ 0 \end{bmatrix}$
  if $\lambda = \bar{p}_1 - 2\beta_1 c \geq 0$ and $\mu_2 = \bar{p}_1 - \bar{p}_2 + c(\alpha - 2\beta_1) \geq 0$
- case 4: $\mathbf{d}^{UNE} = \begin{bmatrix} 0 \\ c \end{bmatrix}$
  if $\lambda = \bar{p}_2 - 2\beta_2 c \geq 0$ and $\mu_1 = \bar{p}_2 - \bar{p}_1 + c(\alpha - 2\beta_2) \geq 0$

$$\begin{aligned}
TS(d) =& d_1(\bar{p}_1 - \frac{1}{2}\beta_1 d_1 - \frac{1}{2}\alpha d_2) + d_2(\bar{p}_2 - \frac{1}{2}\beta_2 d_2 - \frac{1}{2}\alpha d_1) \\
=& d_1\bar{p}_1 + d_2\bar{p}_2 - \frac{1}{2}\left(\beta_1 d_1^2 - \alpha d_2 d_1 + \beta_2 d_2^2\right) \\
& \text{with } d_1 + d_2 = c \\
=& \bar{p}_2 c - \frac{1}{2}\beta_2 c^2 + d_1\left(\bar{p}_1 - \bar{p}_2 - \frac{1}{2}\beta_1 d_1 + \frac{1}{2}\alpha c - \frac{1}{2}\alpha d_1 + \beta_2 c - \frac{1}{2}\beta_2 d_1\right)
\end{aligned}$$

$$\Pi(d) = d_1\bar{p}_1 + d_2\bar{p}_2 - \left(\beta_1 d_1^2 - \alpha d_2 d_1 + \beta_2 d_2^2\right)$$

## B.2 Calculations and proofs for Section 3.4

### B.2.1 Closed-form solutions

Under the assumption of symmetric market, we have $\bar{\mathbf{p}} = \bar{p}_0\mathbf{e}$, and

$$\mathbf{B} = \begin{bmatrix} \beta & \alpha & \dots & \alpha \\ \vdots & \ddots & & \vdots \\ & & \ddots & \\ \alpha & \dots & \alpha & \beta \end{bmatrix}$$

.

The closed-form solutions follow from the KKT conditions presented in Section B.1.1:

$$
\begin{aligned}
d_i^{CP} = & \quad \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{2(\beta+(n-1)\alpha)}\right\} \\
d_i^{SMAX} = & \quad \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{\beta+(n-1)\alpha}\right\} \qquad \forall i = 1, 2, \cdots, n. \\
d_i^{UNE} = & \quad \min\left\{\frac{C}{n}, \frac{\bar{p}_0}{2\beta+(n-1)\alpha}\right\}
\end{aligned}
$$

### B.2.2 Proofs

**Theorem B.3.** *For symmetric subsidiaries facing a single joint capacity constraint, the fraction of profit achieved under free competition compared to the maximum company profit is at least:*

- *When the capacity constraint is not active for both the centrally coordinated problem and the oligopoly problem (i.e.,* $C \geq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}$ *), then:*

$$
\frac{\Pi(OP)}{\Pi(CP)} = \Phi_1(r, n) = \frac{4(1+\frac{1}{\frac{1}{r}-1+\frac{1}{n-1}})}{(2+\frac{1}{\frac{1}{r}-1+\frac{1}{n-1}})^2} \geq \frac{4n}{(n+1)^2} \tag{B.6}
$$

- *When the capacity constraint is active for both problems (i.e.,* $C \leq \frac{\bar{p}_0}{2\alpha+\frac{2(\beta-\alpha)}{n}}$ *), then:*

$$
\frac{\Pi(OP)}{\Pi(CP)} \geq \Phi_2(r, n) = \begin{cases} 1 - \dfrac{\frac{1}{k}-\frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)}\frac{1}{k}+\frac{2}{k}-\frac{1}{n}} > \frac{1}{2}, & \textit{when } \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k+1}} \\ & \textit{for } k = 1, 2, ...n-1 \\ 1 - \dfrac{1-\frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)}+2-\frac{1}{n}} > \frac{1}{2}, & \textit{when } c \leq \frac{\bar{p}_0}{2\beta} \end{cases}
$$

$$
> \frac{1}{2} \tag{B.7}
$$

- *When the constraint is active for the oligopoly problem but inactive for the centrally coordinated problem, the profit ratio lies between the unconstrained bound (B.6) and the constrained bound (B.7).*

*Proof.* When $c \geq \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{n}}$ (constraint is not restrictive for the OP problem),

$$\begin{aligned}
\frac{\Pi(OP)}{\Pi(CP)} =& \frac{n d_i^{OP}(\bar{p}_0 - (\beta + (n-1)\alpha) d_i^{OP})}{n d_i^{CP}(\bar{p}_0 - (\beta + (n-1)\alpha) d_i^{CP})} \\
=& \frac{2(\beta + (n-1)\alpha)}{2\beta + (n-1)\alpha} \frac{(\bar{p}_0 - (\beta + (n-1)\alpha)\frac{\bar{p}_0}{2\beta + (n-1)\alpha})}{\frac{1}{2}\bar{p}_0} \\
=& \frac{4(\beta + (n-1)\alpha)\beta}{(2\beta + (n-1)\alpha)^2} \\
=& \frac{4(1 + \frac{1}{\frac{1}{r} - 1 + \frac{1}{n-1}})}{(2 + \frac{1}{\frac{1}{r} - 1 + \frac{1}{n-1}})^2} \\
>& \frac{4n}{(n+1)^2} (\text{ when } r \to 1)
\end{aligned}$$

When $0 < c < \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{n}}$ (constraint is restrictive for the OP problem),

$$\Pi(OP) = \bar{p}_0 c - \alpha c^2 - (\beta - \alpha) \sum d_i^2$$

- Case 1: $c \leq \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}$ ($\Leftrightarrow \bar{p}_0 \geq 2\beta c$)

$$d_i^w = \begin{cases} c, & i = 1; \\ 0, & \text{otherwise.} \end{cases}$$

$$\begin{aligned}
\Rightarrow \Pi(OP) \geq& \Pi(\mathbf{d}^w) \\
=& \bar{p}_0 c - \alpha c^2 - (\beta - \alpha) c^2
\end{aligned}$$

$$
\begin{aligned}
\frac{\Pi(OP)}{\Pi(CP)} =& \frac{\bar{p}_0 c - \alpha c^2 - (\beta-\alpha)c^2}{\bar{p}_0 c - \alpha c^2 - (\beta-\alpha)\frac{c^2}{n}} \\
=& 1 - \frac{(\beta-\alpha)c^2(1-\frac{1}{n})}{\bar{p}_0 c - \alpha c^2 - (\beta-\alpha)\frac{c^2}{n}} \\
=& 1 - \frac{(\beta-\alpha)c^2(1-\frac{1}{n})}{2\beta c^2 - \alpha c^2 - (\beta-\alpha)\frac{c^2}{n}} \\
=& 1 - \frac{(\beta-\alpha)(1-\frac{1}{n})}{2\beta - \alpha - (\beta-\alpha)\frac{1}{n}} \\
=& 1 - \frac{1-\frac{1}{n}}{\frac{2\beta-\alpha}{\beta-\alpha} - \frac{1}{n}} \\
=& 1 - \frac{1-\frac{1}{n}}{\frac{\alpha}{\beta-\alpha} + 2 - \frac{1}{n}} \\
=& 1 - \frac{1-\frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)} + 2 - \frac{1}{n}}
\end{aligned}
$$

( since $\frac{\alpha}{\beta-\alpha} = \frac{1}{(n-1)(\frac{1}{r}-1)}$)

- Case 2: $\frac{c}{n} \leq \frac{c}{k+1} \leq \frac{\bar{p}_0-\alpha c}{2\beta-\alpha} < \frac{c}{k}$ for some $1 \leq k \leq n-1$.

$$(\Leftrightarrow \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{k+1}})$$

The worst case equilibrium in terms of profit is given by:

$$
d_i^w = \begin{cases} \frac{\bar{p}_0-\alpha c}{2\beta-\alpha} & \forall i = 1, 2, ..., k \\ c - k\frac{\bar{p}_0-\alpha c}{2\beta-\alpha} & i = k+1 \\ 0 & \forall i = k+2, ..., n \end{cases}
$$

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{\bar{p}_0 c - \alpha c^2 - (\beta-\alpha)\sum (d_i^w)^2}{\bar{p}_0 c - \alpha c^2 - \frac{1}{n}(\beta-\alpha)c^2}$$

$$= \frac{\bar{p}_0 c - \alpha c^2 - (\beta-\alpha)(k\left(\frac{\bar{p}_0-\alpha c}{2\beta-\alpha}\right)^2 + \left(c - k\frac{\bar{p}_0-\alpha c}{2\beta-\alpha}\right)^2)}{\bar{p}_0 c - \alpha c^2 - \frac{1}{n}(\beta-\alpha)c^2}$$

$$= 1 - \frac{(k\left(\frac{\bar{p}_0-\alpha c}{2\beta-\alpha}\right)^2 + \left(c - k\frac{\bar{p}_0-\alpha c}{2\beta-\alpha}\right)^2 - \frac{c^2}{n})(\beta-\alpha)}{\bar{p}_0 c - \alpha c^2 - \frac{1}{n}(\beta-\alpha)c^2}$$

Let $x = \frac{\bar{p}_0-\alpha c}{c(2\beta-\alpha)}$, $\bar{p}_0 c = ((2\beta-\alpha)x + \alpha)\, c^2$, $\frac{1}{k+1} \leq x \leq \frac{1}{k}$

$$\frac{\Pi(NE)}{\Pi(CP)} \geq 1 - \frac{(kx^2c^2 + (c - kcx)^2 - \frac{c^2}{n})(\beta-\alpha)}{((2\beta-\alpha)x+\alpha)\, c^2 - \alpha c^2 - \frac{1}{n}(\beta-\alpha)c^2}$$

$$= 1 - \frac{(kx^2 + (1-kx)^2 - \frac{1}{n})(\beta-\alpha)}{(2\beta-\alpha)x - \frac{1}{n}(\beta-\alpha)}$$

$$= 1 - \underbrace{\frac{kx^2 + (1-kx)^2 - \frac{1}{n}}{\frac{2\beta-\alpha}{\beta-\alpha}x - \frac{1}{n}}}_{f(x)}$$

$$f'(\frac{1}{k}) = \frac{2(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n}) - (\frac{1}{k} - \frac{1}{n})\frac{2\beta-\alpha}{\beta-\alpha}}{(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n})^2}$$

$$= \frac{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{2}{n} + \frac{2}{n} + \frac{1}{n}\frac{\alpha}{\beta-\alpha}}{(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n})^2}$$

$$= \frac{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{1}{n}\frac{\alpha}{\beta-\alpha}}{(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n})^2}$$

$$> 0$$

$$
\begin{aligned}
f'(\frac{1}{k+1}) =& \frac{-(\frac{k}{(k+1)^2} + (1-\frac{k}{k+1})^2 - \frac{1}{n})\frac{2\beta-\alpha}{\beta-\alpha}}{(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k+1} - \frac{1}{n})^2} \\
=& \frac{-(\frac{1}{k+1} - \frac{1}{n})\frac{2\beta-\alpha}{\beta-\alpha}}{(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k+1} - \frac{1}{n})^2} \\
<& 0
\end{aligned}
$$

This suggests that $f(x)$ is maximized at the boundary when $x = \frac{1}{k}$ or $\frac{1}{k+1}$. Comparing the two function values:

$$
f(\frac{1}{k}) = \frac{(\frac{1}{k} - \frac{1}{n})}{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n}}
$$

$$
f(\frac{1}{k+1}) = \frac{(\frac{1}{k+1} - \frac{1}{n})}{\underbrace{\frac{2\beta-\alpha}{\beta-\alpha}}_{A}\frac{1}{k+1} - \underbrace{\frac{1}{n}}_{B}}
$$

$$
\begin{aligned}
f(\frac{1}{k}) - f(\frac{1}{k+1}) &= \frac{(\frac{1}{k} - \frac{1}{n})}{A\frac{1}{k} - B} - \frac{(\frac{1}{k+1} - \frac{1}{n})}{A\frac{1}{k+1} - B} \\
&= \frac{(\frac{1}{k} - \frac{1}{n})(A\frac{1}{k+1} - B) - (\frac{1}{k+1} - \frac{1}{n})(A\frac{1}{k} - B)}{(A\frac{1}{k} - B)(A\frac{1}{k+1} - B)} \\
&= \frac{(\frac{1}{k} - \frac{1}{n})(A\frac{1}{k+1} - B) - (\frac{1}{k+1} - \frac{1}{n})(A\frac{1}{k} - B)}{(A\frac{1}{k} - B)(A\frac{1}{k+1} - B)} \\
&= \frac{(\frac{1}{n}A - B)(\frac{1}{k} - \frac{1}{k+1})}{(A\frac{1}{k} - B)(A\frac{1}{k+1} - B)} \\
&> 0
\end{aligned}
$$

Hence $f(x)$ is maximized at $x = \frac{1}{k}$.

$$
\begin{aligned}
\frac{\Pi(OP)}{\Pi(CP)} =&1 - f(x)\\
\geq&1 - f(\frac{1}{k})\\
=&1 - \frac{\frac{1}{k} - \frac{1}{n}}{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} - \frac{1}{n}}\\
=&1 - \frac{\frac{1}{k} - \frac{1}{n}}{\frac{\alpha}{\beta-\alpha}\frac{1}{k} + \frac{2}{k} - \frac{1}{n}}\\
&(\text{ Since } \frac{\alpha}{\beta-\alpha} = \frac{1}{(n-1)(\frac{1}{r}-1)})\\
=&1 - \frac{\frac{1}{k} - \frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)}\frac{1}{k} + \frac{2}{k} - \frac{1}{n}}\\
>&\frac{1}{2-\frac{k}{n}}(\text{ when } r \to 0)
\end{aligned}
$$

$$
\frac{\Pi(OP)}{\Pi(CP)} \geq \begin{cases}
\frac{4(1+\frac{1}{\frac{1}{r}-1+\frac{1}{n-1}})}{(2+\frac{1}{\frac{1}{r}-1+\frac{1}{n-1}})^2} \geq \frac{4n}{(n+1)^2} & \text{when } c \geq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}\\
1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)}\frac{1}{k}+\frac{2}{k}-\frac{1}{n}} > \frac{1}{2}, & \text{when } \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k+1}}\\
& \text{for } k = 1, 2, \ldots n-1\\
1 - \frac{1-\frac{1}{n}}{\frac{1}{(n-1)(\frac{1}{r}-1)}+2-\frac{1}{n}} > \frac{1}{2}, & \text{when } c \leq \frac{\bar{p}_0}{2\beta}
\end{cases}
\qquad \square
$$

**Theorem B.4.** *For symmetric subsidiaries facing a single joint capacity constraint, the fraction of social welfare achieved under free competition compared to the maximum achievable welfare is at least 75%:*

- *When the capacity constraint is not active for both the oligopoly problem and the SMAX problem ($C \geq \frac{\bar{p}_0}{\alpha+\frac{\beta-\alpha}{n}}$), then:*

$$
\frac{TS(OP)}{TS(SMAX)} = \Phi_3(r,n) = \left(\frac{1}{r} + \frac{1}{n-1}\right)\frac{\frac{3}{r}+\frac{3}{n-1}-2}{\left(\frac{2}{r}+\frac{2}{n-1}-1\right)^2} \geq \frac{3}{4}; \qquad \text{(B.8)}
$$

- *When the constraint is active for both problems ($C \leq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}$), then:*

$$\frac{TS(OP)}{TS(SMAX)} \geq \Phi_4(r,n) = \begin{cases} 1-\frac{\frac{1}{k}-\frac{1}{n}}{\frac{4}{k}+\frac{1}{(n-1)(\frac{1}{r}-1)}(\frac{2}{k}+1)-\frac{1}{n}} > \frac{3}{4}, & \text{when } \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k+1}} \\ & \text{for } k=1,2,...n-1 \\ 1-\frac{1}{3(1+\frac{1}{n-1})(1+\frac{1}{(n-1)(\frac{1}{r}-1)})+1} > \frac{3}{4}, & \text{when } c \leq \frac{\bar{p}_0}{2\beta} \end{cases} > \frac{3}{4} \tag{B.9}$$

- *When the constraint is active for the SMAX problem but not for the oligopoly problem the social welfare ratio lies above the unconstrained bound (B.8).*

*Proof.* For $\mathbf{d}$ with $d_1 = d_2 = .... = d_n$,

$$\begin{aligned} TS(\mathbf{d}) &= nd_i(\bar{p}_0 - \frac{1}{2}n\alpha d_i - \frac{1}{2}(\beta-\alpha)d_i) \\ &= nd_i(\bar{p}_0 - \frac{1}{2}(\beta+(n-1)\alpha)d_i) \end{aligned}$$

When $c \geq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}$(constraint is not restrictive for the OP problem),

$$\begin{aligned} \frac{TS(OP)}{TS(SMAX)} &\geq \frac{nd_i^{OP}(\bar{p}_0 - \frac{1}{2}(\beta+(n-1)\alpha)d_i^{OP})}{nd_i^{SMAX}(\bar{p}_0 - \frac{1}{2}(\beta+(n-1)\alpha)d_i^{SMAX})} \\ &= \frac{\beta+(n-1)\alpha}{2\beta+(n-1)\alpha}\frac{1-\frac{1}{2}\frac{\beta+(n-1)\alpha}{2\beta+(n-1)\alpha}}{\frac{1}{2}} \\ &\quad (\text{Since } \frac{1}{r}-1+\frac{1}{n-1} = \frac{\beta}{(n-1)\alpha},) \\ &= \frac{\frac{1}{r}-1+\frac{1}{n-1}+1}{2(\frac{1}{r}-1+\frac{1}{n-1})+1}\frac{1-\frac{1}{2}\frac{\frac{1}{r}-1+\frac{1}{n-1}+1}{2(\frac{1}{r}-1+\frac{1}{n-1})+1}}{\frac{1}{2}} \\ &= \frac{(3(\frac{1}{r}+\frac{1}{n-1})-2)(\frac{1}{r}+\frac{1}{n-1})}{\left[2(\frac{1}{r}+\frac{1}{n-1})-1\right]^2} \geq \frac{3}{4} \end{aligned}$$

The minimum of the ratio $\frac{3}{4}$ is achieved when $r=0$.

When $0 < c < \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}}$ (constraint is restrictive for the OP problem), we have a set of equilibrium strategies satisfying

$$\mathbf{d}^{OP} = \left\{\mathbf{d}|0 \le d_i \le \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}, \sum d_i = c\right\}$$

$$\begin{aligned} TS(SMAX) &= \bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)\sum(\frac{c}{n})^2 \\ &= \bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta - \alpha)c^2. \end{aligned}$$

we have

$$\begin{aligned} TS(OP) =& \sum d_i^{NE}(\bar{p}_0 - \frac{1}{2}(\beta - \alpha)d_i^{NE} - \frac{1}{2}\alpha c) \\ =& \bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)\sum d_i^{NE2} \\ \ge& \bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)\sum d_i^{w2} \end{aligned}$$

Where $\mathbf{d}^w$ is the equilibrium with the largest Euclidean norm.

- Case 1: $c \le \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}$ $(\Leftrightarrow \bar{p}_0 \ge 2\beta c)$

$$d_i^w = \begin{cases} c, & i = 1; \\ 0, & \text{otherwise.} \end{cases}$$

$$\Rightarrow TS(NE) \ge \bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)c^2$$

$$
\begin{aligned}
\frac{TS(OP)}{TS(SMAX)} \geq & \frac{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta-\alpha)c^2}{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta-\alpha)c^2} \\
=& 1 - \frac{\frac{1}{2}(1-\frac{1}{n})(\beta-\alpha)c^2}{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta-\alpha)c^2} \\
& (\text{Since } \bar{p}_0 \geq 2\beta c) \\
\geq & 1 - \frac{\frac{1}{2}(1-\frac{1}{n})(\beta-\alpha)c^2}{(2\beta c)c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta-\alpha)c^2} \\
=& 1 - \frac{\frac{1}{2}(1-\frac{1}{n})(\beta-\alpha)}{2\beta - \frac{1}{2}\alpha - \frac{1}{2n}(\beta-\alpha)} \\
=& 1 - (1-\frac{1}{n})\frac{1}{\frac{4\beta-\alpha}{\beta-\alpha} - \frac{1}{n}} \\
=& 1 - (1-\frac{1}{n})\frac{1}{\frac{3\alpha}{\beta-\alpha} + 4 - \frac{1}{n}} \\
& (\text{Since } \frac{\alpha}{\beta-\alpha} = \frac{1}{(n-1)(\frac{1}{r}-1)}) \\
=& 1 - (1-\frac{1}{n})\frac{1}{\frac{3}{(n-1)(\frac{1}{r}-1)} + 4 - \frac{1}{n}} \\
=& 1 - \frac{1}{3(1+\frac{1}{n-1})(1+\frac{1}{(n-1)(\frac{1}{r}-1)}) + 1} \\
>& 1 - \frac{1}{4+\frac{1}{n-1}}(\text{ when } r \to 1) \\
>& \frac{3}{4}(\text{ when } n \to \infty)
\end{aligned}
$$

- Case 2: $\frac{c}{n} \leq \frac{c}{k+1} \leq \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha} < \frac{c}{k}$ for some $1 \leq k \leq n-1$. ($\Leftrightarrow \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha + \frac{2\beta-\alpha}{k+1}}$)

  The worst case Oligopoly Equilibrium in terms of social welfare is given by:

$$
d_i^w = \begin{cases} \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha} & \forall i = 1, 2, ..., k \\ c - k\frac{\bar{p}_0 - \alpha c}{2\beta - \alpha} & i = k+1 \\ 0 & \forall i = k+2, ..., n \end{cases}
$$

  Because for any $\sum d_i = c$, if there exist index $i$ and $j$ such that $d_i > 0$, $d_j > 0$, and $d_i + d_j \leq \frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}$, we can always construct $\tilde{\mathbf{d}}$ such that $\sum d_i^2 < \sum \tilde{d}_i^2$ by letting

$$\tilde{d}_k = \begin{cases} d_i + d_j & k = i \\ 0 & k = j \\ d_k & \forall k \neq i, k \neq j \end{cases}$$

$$\begin{aligned} \sum d_i^2 - \sum \tilde{d}_i^2 =& d_i^2 + d_j^2 - \tilde{d}_i^2 - \tilde{d}_j^2 \\ =& d_i^2 + d_j^2 - (d_i + d_j)^2 \\ =& - 2d_i d_j \\ <& 0 \end{aligned}$$

Hence

$$\begin{aligned} \frac{TS(OP)}{TS(SMAX)} \geq& \frac{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)\sum (d_i^w)^2}{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta - \alpha)c^2} \\ =& \frac{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2}(\beta - \alpha)(k\left(\frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}\right)^2 + \left(c - k\frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}\right)^2)}{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta - \alpha)c^2} \\ =& 1 - \frac{1}{2}\frac{(k\left(\frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}\right)^2 + \left(c - k\frac{\bar{p}_0 - \alpha c}{2\beta - \alpha}\right)^2 - \frac{c^2}{n})(\beta - \alpha)}{\bar{p}_0 c - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta - \alpha)c^2} \end{aligned}$$

Let $x = \frac{\bar{p}_0 - \alpha c}{c(2\beta - \alpha)}$, we have $\bar{p}_0 c = \left((2\beta - \alpha)x + \alpha\right)c^2$, $\frac{1}{k+1} \leq x < \frac{1}{k}$

$$
\begin{aligned}
\frac{TS(OP)}{TS(SMAX)} \geq & 1 - \frac{1}{2}\frac{(kx^2c^2 + (c-kcx)^2 - \frac{c^2}{n})(\beta-\alpha)}{((2\beta-\alpha)x+\alpha)\,c^2 - \frac{1}{2}\alpha c^2 - \frac{1}{2n}(\beta-\alpha)c^2} \\
= & 1 - \frac{1}{2}\frac{(kx^2 + (1-kx)^2 - \frac{1}{n})(\beta-\alpha)}{(2\beta-\alpha)x + \frac{1}{2}\alpha - \frac{1}{2n}(\beta-\alpha)} \\
= & 1 - \frac{1}{2}\underbrace{\frac{kx^2 + (1-kx)^2 - \frac{1}{n}}{\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}}}_{f(x)}
\end{aligned}
$$

Since

$$
\begin{aligned}
& f'(x) = \\
& \frac{(2kx - 2k\,(1-kx))\,(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}) - \left(kx^2 + (1-kx)^2 - \frac{1}{n}\right)\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
& \frac{2k\,((1+k)x-1)\,(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}) - \left(kx^2 + (1-kx)^2 - \frac{1}{n}\right)\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2}
\end{aligned}
$$

$$
\begin{aligned}
f'(\frac{1}{k}) =& \frac{2(\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}) - \left(\frac{1}{k} - \frac{1}{n}\right)\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
=& \frac{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{\alpha}{\beta-\alpha} - \frac{1}{n} + \frac{1}{n}\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
=& \frac{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{\alpha}{\beta-\alpha} - \frac{1}{n} + \frac{1}{n} + \frac{1}{n}\frac{\beta}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
=& \frac{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{\alpha}{\beta-\alpha} + \frac{1}{n}\frac{\beta}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
>& 0
\end{aligned}
$$

$$
\begin{aligned}
f'(\frac{1}{k+1}) =& \frac{-\left(\frac{k}{(k+1)^2} + (1 - \frac{k}{1+k})^2 - \frac{1}{n}\right)\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
=& \frac{-\left(\frac{1}{1+k} - \frac{1}{n}\right)\frac{2\beta-\alpha}{\beta-\alpha}}{\left(\frac{2\beta-\alpha}{\beta-\alpha}x + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}\right)^2} \\
<& 0
\end{aligned}
$$

This suggests that $f(x)$ is maximized at the boundary when $x = \frac{1}{k}$ or $\frac{1}{k+1}$. Comparing the two function values:

$$
f(\frac{1}{k}) = \frac{(\frac{1}{k} - \frac{1}{n})}{\frac{2\beta-\alpha}{\beta-\alpha}\frac{1}{k} + \frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}}
$$

$$
f(\frac{1}{k+1}) = \frac{(\frac{1}{k+1} - \frac{1}{n})}{\underbrace{\frac{2\beta-\alpha}{\beta-\alpha}}_{A}\frac{1}{k+1} + \underbrace{\frac{1}{2}\frac{\alpha}{\beta-\alpha} - \frac{1}{2n}}_{B}}
$$

$$
\begin{aligned}
f(\frac{1}{k}) - f(\frac{1}{k+1}) &= \frac{(\frac{1}{k}-\frac{1}{n})}{A\frac{1}{k}+B} - \frac{(\frac{1}{k+1}-\frac{1}{n})}{A\frac{1}{k+1}+B} \\
&= \frac{(\frac{1}{k}-\frac{1}{n})(A\frac{1}{k+1}+B) - (\frac{1}{k+1}-\frac{1}{n})(A\frac{1}{k}+B)}{(A\frac{1}{k}+B)(A\frac{1}{k+1}+B)} \\
&= \frac{(\frac{1}{n}A+B)(\frac{1}{k}-\frac{1}{k+1})}{(A\frac{1}{k}+B)(A\frac{1}{k+1}+B)} \\
&> 0
\end{aligned}
$$

Hence $f(x)$ is maximized at $x = \frac{1}{k}$. It is interesting to notice that $f$ is not a monotonically decreasing function in $x$.

$$
\begin{aligned}
\frac{TS(OP)}{TS(SMAX)} &= 1 - \frac{1}{2}f(x) \\
&\geq 1 - \frac{1}{2}f(\frac{1}{k}) \\
&= 1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{2\beta-\alpha}{\beta-\alpha}\frac{2}{k} + \frac{\alpha}{\beta-\alpha} - \frac{1}{n}} \\
&= 1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{4}{k} + \frac{\alpha}{\beta-\alpha}\frac{2}{k} + \frac{\alpha}{\beta-\alpha} - \frac{1}{n}} \\
&= 1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{4}{k} + \frac{\alpha}{\beta-\alpha}(\frac{2}{k}+1) - \frac{1}{n}} \\
&= 1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{4}{k} + \frac{1}{(n-1)(\frac{1}{r}-1)}(\frac{2}{k}+1) - \frac{1}{n}}
\end{aligned}
$$

$$
\frac{TS(OP)}{TS(SMAX)} \geq \begin{cases}
\left(\frac{1}{r}+\frac{1}{n-1}\right)\frac{\frac{3}{r}+\frac{3}{n-1}-2}{\left(\frac{2}{r}+\frac{2}{n-1}-1\right)^2} \geq \frac{3}{4} & \text{when } c \geq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{n}} \\
1 - \frac{\frac{1}{k}-\frac{1}{n}}{\frac{4}{k}+\frac{1}{(n-1)(\frac{1}{r}-1)}(\frac{2}{k}+1)-\frac{1}{n}} > \frac{3}{4}, & \text{when } \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k}} < c \leq \frac{\bar{p}_0}{\alpha+\frac{2\beta-\alpha}{k+1}} \\
 & \text{for } k = 1, 2, \ldots n-1 \\
1 - \frac{1}{3(1+\frac{1}{n-1})(1+\frac{1}{(n-1)(\frac{1}{r}-1)})+1} > \frac{3}{4}, & \text{when } c \leq \frac{\bar{p}_0}{2\beta}
\end{cases}
\qquad \square
$$

## B.3 Calculations and Proofs for Section 3.5

### B.3.1 Closed-form solutions

The closed-form solutions when $\bar{\mathbf{p}} = \bar{p}_0\mathbf{e}$ follows from the KKT conditions presented in Section B.1.1 : $(\bar{\mathbf{d}} = \mathbf{B}^{-1}\bar{\mathbf{p}} = \bar{p}_0\mathbf{B}^{-1}\mathbf{e})$

$$\mathbf{d}^{SMAX} = \min\left\{\bar{p}_0\mathbf{B}^{-1}\mathbf{e}, \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}\mathbf{B}^{-1}\mathbf{e}\right\}$$

$$\mathbf{d}^{CP} = \min\left\{\frac{1}{2}\bar{p}_0\mathbf{B}^{-1}\mathbf{e}, \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}\mathbf{B}^{-1}\mathbf{e}\right\}$$

$$\mathbf{d}^{UNE} = \min\left\{\bar{p}_0(\mathbf{B}+\Gamma)^{-1}\mathbf{e}, \frac{C}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right\}$$

### B.3.2 Proofs

**Lemma B.2.** $(1-r)\Gamma^{-1}\boldsymbol{e} \leq \mathbf{B}^{-1}\boldsymbol{e} \leq \Gamma^{-1}\boldsymbol{e}$

*Proof.* we have

$$\begin{cases} B_{ii}M_{ii} - \sum_{k\neq i} B_{ik}M_{ki} = 1 & \forall i \\ B_{ij}M_{jj} - \sum_{k\neq j} B_{ik}M_{kj} = 0 & \forall j \neq i \end{cases}$$

$$\Rightarrow B_{ii}M_{ii} \geq 1$$

By the definition of the diversion ratio $r$, we have

$$\begin{aligned}
& & rM_{ii} &\geq \textstyle\sum_{j\neq i} M_{ij}\forall i \\
\Rightarrow & & -rM_{ii} &\leq -\textstyle\sum_{j\neq i} M_{ij} \\
\Rightarrow & & (1-r)M_{ii} &\leq M_{ii} - \textstyle\sum_{j\neq i} M_{ij} \\
\Rightarrow & & 1-r \leq (1-r)M_{ii}B_{ii} &\leq \left(M_{ii} - \textstyle\sum_{j\neq i} M_{ij}\right) B_{ii} \\
\Rightarrow & & (1-r)\tfrac{1}{B_{ii}} &\leq M_{ii} - \textstyle\sum_{j\neq i} M_{ij} \\
\Rightarrow & & (1-r)\Gamma^{-1}\mathbf{e} &\leq \mathbf{B}^{-1}\mathbf{e}
\end{aligned}$$

□

**Theorem B.5.** *When the subsidiaries have symmetric price potentials and the capacity constraint is active, the losses of profit and welfare for the worst oligopoly equilibrium are characterized by:*

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{1}{2} + \frac{1}{2}\frac{1}{(2B_{MM}\left(\boldsymbol{e}^T\Gamma^{-1}\boldsymbol{e}\right) - 1)} > \frac{1}{2} \tag{B.10}$$

$$\frac{TS(OP)}{TS(SMAX)} \geq \frac{3}{4} + \frac{3}{4}\frac{1}{(4B_{MM}\boldsymbol{e}^T\Gamma^{-1}\boldsymbol{e} - 1)} > \frac{3}{4} \tag{B.11}$$

*where* $B_{MM} = \max\{B_{ii}\}$.

*Proof.* let $i^M = \text{argmax}_i B_{ii}$, define $\mathbf{d}^w$ as follows:

$$d_k^w = \begin{cases} C & \text{for } k = i^M \\ 0 & \text{otherwise} \end{cases}$$

$\Pi(OP) \geq \Pi(\mathbf{d}^w) = C(\bar{p}_0 - B_{i^M i^M} C)$

$$
\begin{aligned}
\frac{\Pi(OP)}{\Pi(CP)} &\geq \frac{C(\bar{p}_0 - B_{i^M i^M} C)}{C(\bar{p}_0 - \frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}})} \\
&= \frac{\bar{p}_0 - B_{i^M i^M} C}{\bar{p}_0 - \frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= 1 - C \frac{B_{i^M i^M} - \frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{\bar{p}_0 - \frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&\geq 1 - C \frac{\mathbf{B}_{i^M i^M} - \frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{2 B_{i^M i^M} C - \frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= 1 - \frac{B_{i^M i^M} - \frac{1}{\mathbf{e}^T B^{-1} \mathbf{e}}}{2 B_{i^M i^M} - \frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= 1 - \frac{1}{2} + \frac{1}{2} \frac{\frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{2 B_{i^M i^M} - \frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= \frac{1}{2} + \frac{1}{2} \frac{\frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{2 B_{i^M i^M} - \frac{1}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= \frac{1}{2} + \frac{1}{2} \frac{1}{2 B_{i^M i^M} \left(\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}\right) - 1} \\
&\geq \frac{1}{2} + \frac{1}{2} \frac{1}{2 B_{i^M i^M} \left(\mathbf{e}^T \Gamma^{-1} \mathbf{e}\right) - 1} \\
&\geq \frac{1}{2}
\end{aligned}
$$

$TS(OP) \geq TS(\mathbf{d}^w) = C(\bar{p}_0 - \frac{1}{2} B_{i^M i^M} C)$

$$
\begin{aligned}
\frac{TS(OP)}{TS(SMAX)} &\geq \frac{C(\bar{p}_0 - \frac{1}{2}B_{i^M i^M}C)}{C(\bar{p}_0 - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}})} \\
&= \frac{\bar{p}_0 - \frac{1}{2}B_{i^M i^M}C}{\bar{p}_0 - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{C}{2}\frac{B_{i^M i^M} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\bar{p}_0 - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&\geq 1 - \frac{C}{2}\frac{B_{i^M i^M} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{2B_{i^M i^M}C - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{1}{2}\frac{B_{i^M i^M} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{2B_{i^M i^M} - \frac{1}{2}\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{1}{4} + \frac{3}{4}\frac{\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{4B_{i^M i^M} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= \frac{3}{4} + \frac{3}{4}\frac{1}{4B_{i^M i^M}\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e} - 1} \\
&\geq \frac{3}{4} + \frac{3}{4}\frac{1}{4B_{i^M i^M}\mathbf{e}^T\Gamma^{-1}\mathbf{e} - 1} \\
&\geq \frac{3}{4}
\end{aligned}
$$

□

**Theorem B.6.** *When the joint constraint is active and the price potentials are symmetric across subsidiaries, the loss of company profit at the uniform Nash equilibrium compared with the centrally coordinated solution is no more than 1/3:*

$$\frac{\Pi(UNE)}{\Pi(CP)} \geq 2 - 2\delta + \frac{3}{4}\delta^2 \geq \frac{2}{3} \tag{B.12}$$

*where $2 - r \leq \delta \leq 2$. Tightness of the bound is achieved when $r = 0$.*

*Proof.* Let

$$\mathbf{e}^T\mathbf{B}^{-1}\Gamma^{\frac{1}{2}} = \mathbf{x}^T$$

$$\mathbf{e}^T(\mathbf{B} + \Gamma)^{-1}\Gamma^{\frac{1}{2}} = \mathbf{y}^T$$

$$\mathbf{e}^T\Gamma^{-\frac{1}{2}} = \mathbf{z}^T$$

$$
\begin{aligned}
\Rightarrow (\mathbf{x}-\mathbf{y})^T\mathbf{z} =& \mathbf{e}^T\left(\mathbf{B}^{-1}-(\mathbf{B}+\Gamma)^{-1}\right)\Gamma^{\frac{1}{2}}\Gamma^{-\frac{1}{2}}\mathbf{e} \\
=& \mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\left((\mathbf{B}+\Gamma)\mathbf{B}^{-1}-\mathbf{I}\right)\mathbf{e} \\
=& \mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\Gamma\mathbf{B}^{-1}\mathbf{e} \\
=& \mathbf{x}^T\mathbf{y}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\Pi(UNE)}{\Pi(CP)} &= \frac{\bar{p}_0 C - C^2\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{\left(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)^2}}{\bar{p}_0 C - \frac{C^2}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - C\frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{\left(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\bar{p}_0 - \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&\text{(Since } \bar{p}_0 \geq \frac{2C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}\text{)} \\
&\geq 1 - \frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{\left(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 2 - \frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\ \mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}{\left(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)^2} \\
&= 2 + \frac{\left(-\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e} + \mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\Gamma(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}{\left(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}\right)^2} \\
&= 2 + \frac{\left(-\mathbf{y}^T\mathbf{z} + \mathbf{y}^T\mathbf{y}\right)\mathbf{x}^T\mathbf{z}}{(\mathbf{y}^T\mathbf{z})^2} \\
&\geq 2 + \frac{\left(-\mathbf{y}^T\mathbf{z} + \mathbf{x}^T\mathbf{y} - \frac{1}{4}\mathbf{x}^T\mathbf{x}\right)\mathbf{x}^T\mathbf{z}}{(\mathbf{y}^T\mathbf{z})^2} \\
&\geq 2 + \frac{\left(-\mathbf{y}^T\mathbf{z} + \mathbf{x}^T\mathbf{y} - \frac{1}{4}\mathbf{x}^T\mathbf{z}\right)\mathbf{x}^T\mathbf{z}}{(\mathbf{y}^T\mathbf{z})^2} \\
&= 2 + \frac{\left(-\mathbf{y}^T\mathbf{z} + (\mathbf{x}^T\mathbf{z} - \mathbf{y}^T\mathbf{z}) - \frac{1}{4}\mathbf{x}^T\mathbf{z}\right)\mathbf{x}^T\mathbf{z}}{(\mathbf{y}^T\mathbf{z})^2} \\
&= 2 + \frac{\left(-2\mathbf{y}^T\mathbf{z} + \frac{3}{4}\mathbf{x}^T\mathbf{z}\right)\mathbf{x}^T\mathbf{z}}{(\mathbf{y}^T\mathbf{z})^2} \\
&= 2 - 2\delta + \frac{3}{4}\delta^2
\end{aligned}
$$

By Lemma B.2, we have$(1-r)\mathbf{z} \le \mathbf{x} \le \mathbf{z}$.

$$\begin{aligned}
\Rightarrow \tilde{\delta} &= \frac{\mathbf{x}^T\mathbf{z}}{\mathbf{y}^T\mathbf{z}} \\
&= \frac{\mathbf{x}^T\mathbf{y} + \mathbf{y}^T\mathbf{z}}{\mathbf{y}^T\mathbf{z}} \\
&= 1 + \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{y}^T\mathbf{z}} \\
&\ge 1 + (1-r)\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{y}^T\mathbf{z}} \\
&\ge 2 - r
\end{aligned}$$

Hence, $2 - r \le \delta \le 2$. So $\frac{\Pi(UNE)}{\Pi(CP)}$ is minimized when $\delta = \frac{4}{3}$ with a minimum of $\frac{2}{3}$. □

**Theorem B.7.** *When the joint constraint is active and the price potentials are symmetric across subsidiaries, the loss of social welfare at the uniform Nash equilibrium compared with the centrally coordinated solution is no more than 1/4:*

$$\frac{TS(UNE)}{TS(SMAX)} \ge \max\left\{\frac{2}{3} + \frac{2}{3\,(2 + r(n-1))}, \frac{3}{8}\frac{(2-r)^2}{\frac{3}{2} - r}\right\} \ge \frac{3}{4}. \tag{B.13}$$

*The first bound dominates when $n$ is small while the second bound dominates when $n$ is large.*

*Proof.*

$$\begin{aligned}
\frac{TS(UNE)}{TS(SMAX)} &= \frac{\bar{p}_0 C - \frac{C^2}{2}\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2}}{\bar{p}_0 C - \frac{1}{2}\frac{C^2}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{C}{2}\frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\bar{p}_0 - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&\text{(Since } \bar{p}_0 \geq \frac{C}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}\text{)} \\
&\geq 1 - \frac{C}{2}\frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{B}(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{C}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}} - \frac{1}{2}\frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{1}{2}\frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e} - \mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\Gamma(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{1}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}} - \frac{1}{2}\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= 1 - \frac{1}{2}\frac{\frac{1}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}} - \frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\Gamma(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2} - \frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{1}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}} - \frac{1}{2}\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= \frac{1}{2} + \frac{1}{2}\frac{\frac{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\Gamma(\mathbf{B}+\Gamma)^{-1}\mathbf{e}}{(\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e})^2} + \frac{1}{2}\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{1}{\mathbf{e}^T(\mathbf{B}+\Gamma)^{-1}\mathbf{e}} - \frac{1}{2}\frac{1}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= \frac{1}{2} + \frac{1}{2}\frac{\frac{\mathbf{y}^T\mathbf{y}}{\mathbf{y}^T\mathbf{z}} + \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}}{1 - \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}} \\
&\geq \frac{1}{2} + \frac{1}{2}\frac{\frac{\mathbf{x}^T\mathbf{y} - \frac{1}{4}\mathbf{x}^T\mathbf{x}}{\mathbf{y}^T\mathbf{z}} + \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}}{1 - \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}} \\
&= \frac{1}{2} + \frac{1}{2}\frac{\frac{\mathbf{x}^T\mathbf{z} - \mathbf{y}^T\mathbf{z} - \frac{1}{4}\mathbf{x}^T\mathbf{z}}{\mathbf{y}^T\mathbf{z}} + \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}}{1 - \frac{1}{2}\frac{\mathbf{y}^T\mathbf{z}}{\mathbf{x}^T\mathbf{z}}}
\end{aligned}$$

Let $\tilde{\delta} = \frac{\mathbf{x}^T\mathbf{z}}{\mathbf{y}^T\mathbf{z}}$.

$$\begin{aligned}
\frac{TS(UNE)}{TS(SMAX)} &= \frac{1}{2} + \frac{1}{2}\frac{(1 - \frac{1}{4})\tilde{\delta} - 1 + \frac{1}{2}\frac{1}{\tilde{\delta}}}{1 - \frac{1}{2}\frac{1}{\tilde{\delta}}} \\
&= \frac{3}{8}\underbrace{\frac{\tilde{\delta}^2}{\tilde{\delta} - \frac{1}{2}}}_{f(\tilde{\delta})}
\end{aligned}$$

$f'(\tilde{\delta}) = \frac{2\tilde{\delta}(\tilde{\delta}-\frac{1}{2})-\tilde{\delta}^2}{(\tilde{\delta}-\frac{1}{2})^2} = \frac{\tilde{\delta}(\tilde{\delta}-1)}{(\tilde{\delta}-\frac{1}{2})^2} \geq 0$ since $\tilde{\delta} \in [1,2]$

Hence $\frac{TS(UNE)}{TS(SMAX)} \geq \frac{3}{8}\frac{(2-r)^2}{\frac{3}{2}-r} \geq \frac{3}{4}$. The second bound $\frac{2}{3} + \frac{2}{3(2+r(n-1))}$ follows from $\|d\|_B^2 \leq (1+r(n-1))\,\|d\|_\Gamma^2$ .(See Kluberg and Perakis [76]) $\square$

## B.4 Calculations and Proofs for Section 3.6

**Theorem B.8.** *When the capacity constraint is active, the losses of profit and welfare for the worst oligopoly equilibrium are characterized by:*

$$\frac{\Pi(OP)}{\Pi(CP)} \geq \frac{\bar{p}_{min} - B_{MM}C}{\bar{p}_{max}} > \theta - \frac{1}{2} \qquad \textit{for} \quad \bar{p}_{max} \leq 2\bar{p}_{min} \tag{B.14}$$

$$\frac{TS(OP)}{TS(SMAX)} \geq \frac{\bar{p}_{min} - 1/2B_{MM}C}{\bar{p}_{max}} > \theta - \frac{1}{4} \qquad \textit{for} \quad \bar{p}_{max} \leq 4\bar{p}_{min} \tag{B.15}$$

*where* $B_{MM} = \max\{B_{ii}\}$.

*Proof.* Let $i^M = \text{argmax}_i B_{ii}$, and let $\mathbf{d}_{op}^w$ be an equilibrium solution that achieves the worst oligopoly profit $\Pi$. We can bound $\Pi(OP)$ and $\Pi(CP)$ by:

$$\begin{aligned}
\Pi(OP) \geq \Pi(\mathbf{d}_{op}^w) &= (\mathbf{d}_{op}^w)^T(\bar{p} - \mathbf{B}\ \mathbf{d}_{op}^w) \qquad \text{with} \sum_i d_i^w = C \\
&\geq (\mathbf{d}_{op}^w)^T(\bar{p}_{min}\ \mathbf{e} - \mathbf{B}\ \mathbf{d}_{op}^w) \\
&\geq C(\bar{p}_{min} - \mathbf{B}_{i^M i^M} C)
\end{aligned}$$

$$\begin{aligned}
\Pi(CP) &= \max_{\mathbf{d}} \left\{\mathbf{d}^T(\bar{\mathbf{p}} - \mathbf{B}\ \mathbf{d}) \quad | \quad \mathbf{e}^T\ \mathbf{d} \leq C, \mathbf{d} \geq \mathbf{0}\right\} \\
&\leq \max_{\mathbf{d}} \left\{\mathbf{d}^T(\bar{p}_{max}\ \mathbf{e} - \mathbf{B}\ \mathbf{d}) \quad | \quad \mathbf{e}^T\ \mathbf{d} \leq C, \mathbf{d} \geq \mathbf{0}\right\} \\
&= C\ (\bar{p}_{max} - \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}})
\end{aligned}$$

$$
\begin{aligned}
\frac{\Pi(OP)}{\Pi(CP)} &\geq \frac{C(\bar{p}_{min} - B_{i^M i^M} C)}{C(\bar{p}_{max} - \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}})} \\
&= \frac{\bar{p}_{min} - B_{i^M i^M} C}{\bar{p}_{max} - \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}} \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ \frac{1 - \frac{B_{i^M i^M}}{\bar{p}_{min}} C}{1 - \frac{1/\bar{p}_{max}}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}} C} \right] = \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - C \frac{\frac{B_{i^M i^M}}{\bar{p}_{min}} - \frac{1/\bar{p}_{max}}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{1 - \frac{1/\bar{p}_{max}}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}} C} \right] \\
&\geq \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - C \frac{\frac{B_{i^M i^M}}{\bar{p}_{min}} - \frac{1/\bar{p}_{max}}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}}{\frac{2\ B_{i^M i^M}\ C}{\bar{p}_{max}} - \frac{1/\bar{p}_{max}}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}} C} \right] \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - \frac{\bar{p}_{max}}{2\bar{p}_{min}} + \frac{\frac{1}{\bar{p}_{max}} - \frac{1}{2\ \bar{p}_{min}}}{\frac{2\ B_{i^M i^M}\ \mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}}{\bar{p}_{max}} - \frac{1}{\bar{p}_{max}}} \right] \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{2} + \frac{\frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{2}}{\underbrace{2\ B_{i^M i^M}\ \mathbf{e}^T\mathbf{B}^{-1}\mathbf{e} - 1}_{\geq 0}} \geq \frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{2} \qquad \forall \bar{p}_{max} \leq 2\bar{p}_{min}
\end{aligned}
$$

Let $\mathbf{d}_{ts}^{w}$ be an equilibrium solution that achieves the worst social welfare $TS$. We can bound $TS(OP)$ and $TS(SMAX)$ by:

$$
\begin{aligned}
TS(OP) \geq TS(\mathbf{d}_{ts}^{w}) &= (\mathbf{d}_{ts}^{w})^T(\bar{p} - 1/2\mathbf{B}\ \mathbf{d}_{ts}^{w}) \qquad \text{with } \sum_i d_i^w = C \\
&\geq (\mathbf{d}_{ts}^{w})^T(\bar{p}_{min}\ \mathbf{e} - 1/2\mathbf{B}\ \mathbf{d}_{ts}^{w}) \\
&\geq C(\bar{p}_{min} - \frac{1}{2} B_{i^M i^M} C)
\end{aligned}
$$

$$
\begin{aligned}
TS(SMAX) &= \max_{\mathbf{d}} \left\{ \mathbf{d}^T(\bar{\mathbf{p}} - 1/2\ \mathbf{B}\ \mathbf{d}) \quad | \quad \mathbf{e}^T\ \mathbf{d} \leq C, \mathbf{d} \geq \mathbf{0} \right\} \\
&\leq \max_{\mathbf{d}} \left\{ \mathbf{d}^T(\bar{p}_{max}\ \mathbf{e} - 1/2\ \mathbf{B}\ \mathbf{d}) \quad | \quad \mathbf{e}^T\ \mathbf{d} \leq C, \mathbf{d} \geq \mathbf{0} \right\} \\
&= C\ (\bar{p}_{max} - \frac{1}{2} \frac{C}{\mathbf{e}^T\mathbf{B}^{-1}\mathbf{e}})
\end{aligned}
$$

$$
\begin{aligned}
\frac{TS(OP)}{TS(SMAX)} &\geq \frac{C(\bar{p}_{min} - \frac{1}{2} B_{i^M i^M} C)}{C(\bar{p}_{max} - \frac{1}{2}\frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}})} = \frac{\bar{p}_{min} - \frac{1}{2} B_{i^M i^M} C}{\bar{p}_{max} - \frac{1}{2}\frac{C}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}} \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ \frac{1 - 1/2 \ \frac{B_{i^M i^M}}{\bar{p}_{min}} C}{1 - 1/2 \ \frac{1/\bar{p}_{max}}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}} C} \right] = \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - \frac{C}{2} \ \frac{\frac{B_{i^M i^M}}{\bar{p}_{min}} - \frac{1/\bar{p}_{max}}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{1 - \frac{1}{2} \ \frac{1/\bar{p}_{max}}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}} C} \right] \\
&\geq \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - \frac{C}{2} \ \frac{\frac{B_{i^M i^M}}{\bar{p}_{min}} - \frac{1/\bar{p}_{max}}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}}}{\frac{2 \ B_{i^M i^M} \ C}{\bar{p}_{max}} - \frac{1}{2} \ \frac{1/\bar{p}_{max}}{\mathbf{e}^T \mathbf{B}^{-1} \mathbf{e}} C} \right] \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} \left[ 1 - \frac{\bar{p}_{max}}{4 \bar{p}_{min}} + 1/2 \ \frac{1 - \frac{\bar{p}_{max}}{4 \ \bar{p}_{min}}}{2 \ B_{i^M i^M} \ \mathbf{e}^T \mathbf{B}^{-1} \mathbf{e} - \frac{1}{2}} \right] \\
&= \frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{4} + \underbrace{\frac{\frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{4}}{4 \ B_{i^M i^M} \ \mathbf{e}^T \mathbf{B}^{-1} \mathbf{e} - 1}}_{\geq 0} \geq \frac{\bar{p}_{min}}{\bar{p}_{max}} - \frac{1}{4} \qquad \forall \bar{p}_{max} \leq 4 \bar{p}_{min}
\end{aligned}
$$

$\square$

**Theorem B.9.** *When the constraint is active for the oligopoly problem, the loss of profit and social welfare under the uniform Nash equilibrium are bounded by:*

$$
\begin{aligned}
TS(UNE) &\geq \tfrac{5}{6} TS(CP) \\
TS(CP) &\geq \tfrac{2}{3} TS(SMAX) \\
TS(UNE) &\geq \tfrac{5}{9} TS(SMAX)
\end{aligned}
$$

*Proof.* (Theorem B.9)

At a Nash equilibrium solution, the optimization problem that a single subsidiary faces is:

$$
\max_{d_i} \ d_i \cdot \left\{ \bar{\mathbf{p}}_i - B_i \cdot \begin{pmatrix} d_i \\ \mathbf{d}_{-i}^{UNE} \end{pmatrix} \right\}
$$

$$
\text{s.t.} \qquad \mathbf{d}_i \in K_i
$$

Let's denote by $\Gamma$ the block diagonal matrix:

$$\Gamma = \begin{pmatrix} B_{11} & 0 & \dots & 0 \\ 0 & B_{22} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & B_{nn} \end{pmatrix}$$

The variational inequality[1] satisfied at the uniform Nash equilibrium is:

$$\{-\bar{\mathbf{p}} + \mathbf{B}\cdot\mathbf{d}^{UNE} + \Gamma\cdot\mathbf{d}^{UNE}\}^T(\mathbf{d} - \mathbf{d}^{UNE}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Since by definition the centrally coordinated solution must be feasible as well (i.e. $\mathbf{d}^{CP} \in K$), we have:

$$\{-\bar{\mathbf{p}} + \mathbf{B}\cdot\mathbf{d}^{UNE} + \Gamma\cdot\mathbf{d}^{UNE}\}^T(\mathbf{d}^{CP} - \mathbf{d}^{UNE}) \geq 0$$

Developing all the terms and extracting $TS(UNE) = (\mathbf{B}\bar{\mathbf{d}})^T\ \mathbf{d}^{UNE} - 1/2\ (\mathbf{d}^{UNE})^T\mathbf{B}\ \mathbf{d}^{UNE}$ and $TS(CP) = (\mathbf{B}\bar{\mathbf{d}})^T\ \mathbf{d}^{CP} - 1/2\ (\mathbf{d}^{CP})^T\mathbf{B}\ \mathbf{d}^{CP}$, we obtain:

$$\begin{aligned} TS(UNE) - TS(CP) - 1/2\ (\mathbf{d}^{UNE})^T\mathbf{B}\mathbf{d}^{UNE} - (\mathbf{d}^{UNE})^T\Gamma\mathbf{d}^{UNE} \\ -1/2\ (\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{CP} + (\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{UNE} + (\mathbf{d}^{CP})^T\Gamma\mathbf{d}^{UNE} \geq 0 \end{aligned} \qquad \text{(B.16)}$$

Using positive definiteness of matrix $\mathbf{B}$, we have:

- $(\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{UNE} \leq 1/2(\mathbf{d}^{UNE})^T\mathbf{B}\mathbf{d}^{UNE} + 1/2(\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{CP}$
  $(1/2\ (\mathbf{d}^{UNE} - \mathbf{d}^{CP})^T\ \mathbf{B}(\mathbf{d}^{UNE} - \mathbf{d}^{CP}) \geq 0)$

- $(\mathbf{d}^{CP})^T\Gamma\mathbf{d}^{UNE} \leq (\mathbf{d}^{UNE})^T\Gamma\mathbf{d}^{UNE} + 1/4(\mathbf{d}^{CP})^T\Gamma\mathbf{d}^{CP}$
  $((\mathbf{d}^{UNE} - \frac{1}{2}\mathbf{d}^{CP})^T\ \Gamma(\mathbf{d}^{UNE} - \frac{1}{2}\mathbf{d}^{CP}) \geq 0)$

[1] A variational inequality VI(F,K) is the problem of finding: $\{x^* \in K,\ F(x^*)^T(x - x^*) \geq 0,\ \forall x \in K\}$.

Hence, using these two inequalities in (B.16), we get:

$$TS(UNE) - TS(CP) + 1/4(\mathbf{d}^{CP})^T \Gamma \mathbf{d}^{CP} \geq 0 \tag{B.17}$$

On the other hand, the variational inequality for the centrally coordinated solution is:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{CP} + \mathbf{B} \cdot \mathbf{d}^{CP}\}^T (\mathbf{d} - \mathbf{d}^{CP}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Evaluating this variational inequality at the feasible vector $0 \in K$ , we have:

$$\bar{\mathbf{p}} \cdot \mathbf{d}^{CP} - 1/2\,(\mathbf{d}^{CP})^T \mathbf{B} \mathbf{d}^{CP} - 3/2\,(\mathbf{d}^{CP})^T \mathbf{B} \mathbf{d}^{CP} \geq 0$$

This leads to:

$$TS(CP) \geq 3/2\,(\mathbf{d}^{CP})^T \mathbf{B} \mathbf{d}^{CP} \Leftrightarrow 1/6\; TS(CP) \geq 1/4\,(\mathbf{d}^{CP})^T \mathbf{B} \mathbf{d}^{CP}$$

For gross substitute products, matrix $\mathbf{B}$ is an inverse M-matrix with all coefficients non negative so that: $(\mathbf{d}^{CP})^T \Gamma \mathbf{d}^{CP} \leq (\mathbf{d}^{CP})^T \mathbf{B} \mathbf{d}^{CP}$. Using (B.18) and this inequality, (B.17) becomes:

$$\begin{aligned} & TS(UNE) - TS(CP) + 1/6\,TS(CP) \geq 0 \\ \Rightarrow\;\; & TS(UNE) \geq 5/6\; TS(CP) \end{aligned}$$

This establishes the desired relation between TS(UNE) and TS(CP). Let us now turn to the study of $TS(SMAX)$. The optimization of the social welfare corresponds to the problem:

$$\max_{\mathbf{d}}\; \bar{\mathbf{p}} \cdot \mathbf{d} - 1/2\; \mathbf{d}^T \mathbf{B}\; \mathbf{d}$$

The optimal solution of this problem satisfies the variational inequality:

$$\{-\bar{\mathbf{p}} + \mathbf{B} \cdot \mathbf{d}^{SMAX}\}^T (\mathbf{d} - \mathbf{d}^{SMAX}) \geq 0 \qquad \forall \mathbf{d} \in K$$

Evaluating this variational inequality at the feasible point 0, we have:

$$TS(SMAX) \geq 1/2\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{SMAX} \tag{B.18}$$

On the other hand, the centrally coordinated variational inequality evaluated at the feasible solution $\mathbf{d}^{SMAX}$ gives:

$$\{-\bar{\mathbf{p}} + \mathbf{B}\cdot\mathbf{d}^{CP} + \mathbf{B}\cdot\mathbf{d}^{CP}\}^T\,(\mathbf{d}^{SMAX} - \mathbf{d}^{CP}) \geq 0$$

This leads to:

$$TS(CP) - 3/2\,(\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{CP} - TS(SMAX) - 1/2\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{SMAX} + 2\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{CP} \geq 0 \tag{B.19}$$

By positive definiteness of matrix $\mathbf{B}$, we have:

$$2\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{CP} \leq 3/2\,(\mathbf{d}^{CP})^T\mathbf{B}\mathbf{d}^{CP} + 2/3\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{SMAX}$$

Making use of this property in (B.19), we finally obtain:

$$\begin{aligned}
& TS(CP) - TS(SMAX) + 1/6\,(\mathbf{d}^{SMAX})^T\mathbf{B}\mathbf{d}^{SMAX} \geq 0 \\
\Rightarrow\ & TS(CP) - TS(SMAX) + 1/3\,TS(SMAX) \geq 0 \quad \text{(by (B.18))} \\
\Rightarrow\ & TS(CP) \geq 2/3\,TS(SMAX)
\end{aligned}$$

Finally, combining $TS(UNE) \geq 5/6\,TS(CP)$ and $TS(CP) \geq 2/3\,TS(SMAX)$, we obtain the desired property:

$$TS(UNE) \geq 5/9\,TS(SMAX)$$

□

# Appendix C

# Proofs for Chapter 4

## C.1 Derivations for unconstrained symmetric generators

### C.1.1 Supply function equilibrium derivation

Under the symmetry assumption, the generators all experience the same quadratic production costs: $C(q) = bq + 1/2\ c\ q^2$. The generators place bids to the system operator by submitting affine supply functions: $q_i(p) = \beta_i(p - \alpha_i),\ i = 1, \ldots, n$. In order to choose its bid, generator $i$ attempts to maximize its profit (taking the supply bids of the competitors as given) for all possible realizations of the demand uncertainty $\epsilon$:

$$\forall \epsilon \geq 0, \qquad p^*_\epsilon,\ q^*_{i,\epsilon} = \begin{array}{rl} \operatorname*{argmax}_{p,\ q} & p\, q - C(q) \\ \text{s.t.} & q + \sum_{j \neq i} q_j(p) = D(p, \epsilon), \end{array}$$

Hence, for every realization of $\epsilon$, generator $i$ chooses a point $(p^*_\epsilon, q^*_{i,\epsilon})$ that maximizes the objective $p[D(p,\epsilon) - \sum_{j \neq i} q_j(p)] - C\left(D(p,\epsilon) - \sum_{j \neq i} q_j(p)\right)$. Setting the derivative (with respect to $p$) to zero leads to :

$$q_{i,\epsilon}^* = [p_\epsilon^* - C'(q_{i,\epsilon}^*)]\left(-\frac{dD}{dp} + \sum_{j\neq i}\frac{dq_j}{dp}\right) \tag{C.1}$$

Generator $i$ chooses the pair $(\alpha_i, \beta_i)$ of its supply bid before knowing the realization of demand. The bid forces generator $i$ to produce $q_{i,\epsilon} = \beta_i(p_\epsilon^* - \alpha_i)$ while optimality condition (C.1) must be satisfied for all realizations of $\epsilon$. The two conditions can be combined into:

$$\beta_i(p_\epsilon^* - \alpha_i) = [p_\epsilon^* - b - c\beta_i(p_\epsilon^* - \alpha_i)]\left(m + \sum_{j\neq i}\beta_j\right) \qquad \forall\epsilon \geq 0 \tag{C.2}$$

Since supply must match demand $\sum_{j=1}^n \beta_j\ p_\epsilon^* - \sum_{j=1}^n \beta_j\alpha_j = -mp_\epsilon^* + \epsilon$, it is clear that the equilibrium price $p_\epsilon^*$ varies with the uncertainty $\epsilon$. The only way optimality condition (C.2) can be satisfied $\forall\epsilon$, is that generator $i$ chooses the pair $(\alpha_i, \beta_i)$ so that the polynomials (in $p_\epsilon^*$) on both sides of the equation are equal. We must have:

$$\beta_i = (1 - c\beta_i)(m + \sum_{j\neq i}\beta_j) \tag{C.3}$$

$$\alpha_i\beta_i = (b - c\beta_i\alpha_i)(m + \sum_{j\neq i}\beta_j) \tag{C.4}$$

After simplification (multiplying (C.3) by $\alpha_i$ and subtracting (C.4)), the linear supply function equilibrium is thus characterized by the system of equations:

$$\begin{cases} \alpha_i = b \\ \beta_i = (1 - c\beta_i)(m + \sum_{j\neq i}\beta_j) \end{cases} \qquad i = 1, \ldots, n$$

As shown in [75], this system of equations has a unique solution vector $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ and this solution is symmetric, meaning $\alpha_1 = \cdots = \alpha_n = b$ and $\beta_1 = \cdots = \beta_n = \beta$. As stated in the beginning of Section 4.3, the supply function bidding mechanism is partially truth revealing as generators end up bidding the true linear part of their production costs $b$ as the intercept $\alpha$ of their supply function bid: $\alpha = b$.

Under our symmetry assumption, there is no loss of generality in setting $b = 0$ because, since $\alpha_i = b$ for all $i$, this linear component of the generators' costs gets directly reflected in the bids, in the production quantities and in all the relevant measures of interest.

A monopolistic generator hence chooses $\beta$ such that $\beta = (1 - c\beta)\ m$ or:

$$\beta = \frac{m}{1 + c\ m} \tag{C.5}$$

.

Under the realization $\epsilon$ of the demand uncertainty, it charges a price $p_\epsilon^*$ such that $\beta\ p_\epsilon^* = -m\ p_\epsilon^* + \epsilon$ and produces a quantity $q_\epsilon^* = \beta\ p_\epsilon^*$. This results in:

$$p_\epsilon^* = \frac{\epsilon}{m}\,\frac{1}{1 + \frac{1}{1+mc}} \qquad \text{and} \qquad q_\epsilon^* = \frac{\epsilon}{2 + cm} \tag{C.6}$$

Similarly, multiple symmetric generators choose $\beta$ such that $\beta = (1-c\beta)[m+(n-1)\beta]$. Solving this quadratic equation results in:

$$\beta = \frac{-m + \frac{n-2}{c} + \sqrt{(-m + \frac{n-2}{c})^2 + \frac{4m(n-1)}{c}}}{2\ (n-1)} \tag{C.7}$$

Under the realization $\epsilon$ of the demand uncertainty, they charge a price $p_\epsilon^*$ such that $n\ \beta\ p_\epsilon^* = -m\ p_\epsilon^* + \epsilon$ and produce a quantity $q_\epsilon^* = \beta\ p_\epsilon^*$. This leads to:

$$p_\epsilon^* = \frac{\epsilon}{\frac{n-2}{2(n-1)}m + \frac{n(n-2)}{2c(n-1)} + \frac{n}{2(n-1)}\sqrt{(-m + \frac{n-2}{c})^2 + \frac{4m(n-1)}{c}}} \quad \text{and} \quad q_\epsilon^* = \frac{\epsilon}{n + \frac{m}{\beta}} \tag{C.8}$$

It is easy to check that $\widetilde{\beta} \triangleq \frac{\beta}{m}$ (in (C.5) and (C.7)) only depends on the parameter $m$ of the demand and the parameter $c$ of the production costs through $z = mc$. This result will be useful in the next subsection.

### C.1.2 Calculations of surplus, profit and welfare ratios

The first step in evaluating the performance ratios between the deregulated and the regulated settings is to describe the regulated solution. As explained in Section 4.2, under a centrally regulated setting, the system operator dispatches the generators to maximize social welfare by solving the optimization:

$$
\begin{aligned}
(\widetilde{p}_\epsilon, \widetilde{\mathbf{q}}_\epsilon) \quad = \quad & \underset{p,\mathbf{q}}{\text{argmax}} \quad TS_\epsilon(p, \mathbf{q}) \\
& \text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} q_i = D(p, \epsilon) \\ 0 \leq q_i \\ \Phi q \leq \Delta \end{cases}
\end{aligned}
$$

where $TS_\epsilon(p, \mathbf{q})$ is the sum of the generators profit and the consumer surplus (see (4.6)).

It can easily be shown that in the absence of constraints, the optimal solution for society is to force the generators to produce enough so that electricity price equals the generators marginal production costs. Under our symmetry assumption, the generators experience the same quadratic production costs: $C(q) = bq + 1/2\ c\ q^2$ and we explained above that, without loss of generality, we take $b = 0$. The optimality condition can thus be translated into $C'(q) = cq = \widetilde{p}$. Matching supply with demand, this leads to:

$$
\begin{gathered}
(C')^{-1}(\widetilde{p}) = -m\widetilde{p} + \epsilon \\
\Rightarrow \widetilde{p} = \frac{\epsilon}{m + \frac{n}{c}} \qquad \text{and} \qquad \widetilde{q} = \frac{\epsilon}{\frac{mc}{n} + 1}
\end{gathered}
\tag{C.9}
$$

It is easy to show that $\beta$ given in (C.7) is always smaller than the inverse of $c$:

$$\beta \leq \frac{1}{c}$$

This results in $p(Dereg.) = p^* \geq p(Centr.) = \widetilde{p}$ and in $q(Dereg.) \leq q(Centr.)$.

We now express the measures of interest (surplus, profit and welfare) as functions of the price and the production quantity of each generator. We will then plug in these expressions the price and quantity of the deregulated and the centralized settings to evaluate the performance ratios.

As defined in (4.5), the consumer surplus is the difference between the utility consumers derive from using electricity and the price they pay for it: $CS_\epsilon(p,d) = U_\epsilon(d) - p\,d$. Removing $d$ from the equation through the demand relationship $D(p,\epsilon) = -m\,p + \epsilon$, we have:

$$CS_\epsilon(p) = \frac{\epsilon^2}{2m} + p\left(\frac{m\,p}{2} - \epsilon\right) \tag{C.10}$$

The generators profit on the other hand can simply be expressed as:

$$\text{Profit}(p,q) = n(p\,q - C(q)) \tag{C.11}$$

Finally, the social welfare is the sum of the generators profit and the consumer surplus:

$$\begin{aligned} TS_\epsilon(p,q) &= CS_\epsilon(p) + n(p\,q - C(q)) \\ &= U_\epsilon\Big(nq\Big) - n\,C(q) \end{aligned} \tag{C.12}$$

The ratios of interest can now be computed. The consumer surplus ratios is:

$$\frac{CS(Dereg.)}{CS(Centr.)} = \frac{CS(p^*_\epsilon)}{CS(\widetilde{p})} = \frac{1 - \frac{1}{1+\frac{n\beta}{m}}}{1 - \frac{1}{1+\frac{n}{mc}}} = \frac{1 - \frac{1}{1+n\widetilde{\beta}}}{1 - \frac{1}{1+\frac{n}{z}}}$$

where $z = mc$ and $p^*_\epsilon$, $\widetilde{p}$ and $\beta$ are defined in (C.6), (C.9) and (C.7) respectively.

Similarly, the ratio of generators profit under the deregulated and the regulated settings is:

$$\frac{\text{Profit}(Dereg.)}{\text{Profit}(Centr.)} = \frac{\text{Profit}(p^*_\epsilon, q^*_\epsilon)}{\text{Profit}(\widetilde{p},\widetilde{q})} = \frac{\left(1 - \frac{c}{2}\beta\right)\frac{1}{(1+\frac{n\beta}{m})^2}\frac{n\beta}{m}}{\frac{1}{2}\frac{1}{(1+\frac{n}{mc})^2}\frac{n}{mc}} = \frac{(1-\frac{z}{2}\widetilde{\beta})\frac{1}{(1+n\widetilde{\beta})^2}n\widetilde{\beta}}{\frac{1}{2}\frac{1}{(1+\frac{n}{z})^2}\frac{n}{z}}$$

Finally, the ratio of social welfare can be expressed as:

$$\frac{TS(Dereg.)}{TS(Centr.)} = \frac{CS(Dereg.) + \text{Profit}(Dereg.)}{CS(Centr.) + \text{Profit}(Centr.)}$$

It is seen from the expressions above that all these performance ratios are independent of the realization of demand uncertainty $\epsilon$ and they depend on the market parameters $m$ and $c$ only through the unit-less parameter $z = mc$.

This allows us to plot the curves of Section 4.2 and to derive the corresponding bounds by carrying some analysis to find the minimum of these ratios as functions of $z$. The profit ratio is a decreasing function of $z$ that tends to 1 at infinity. Therefore the profit ratio always stays above 1: the generators are better off competing than under central coordination. The consumer surplus and social welfare ratios, on the other hand, are both unimodal and they achieve their minimum for the same value of $z = -n + 2\sqrt{n^2 - n}$.

## C.2 Derivations for unconstrained asymmetric generators

Let's first establish the expressions of the consumer surplus and the generators profit for the case of unconstrained asymmetric generators. As established in Appendix C.1, the surplus of consumers can be expressed as:

$$CS_\epsilon(p) = \frac{\epsilon^2}{2m} + p\left(\frac{m\,p}{2} - \epsilon\right)$$

Using the electricity price of the deregulated setting (4.13) $p^* = \frac{\epsilon/m}{1+\sum_{i=1}^n \widetilde{\beta}_i}$ and the regulated setting (4.4.2) $\widetilde{p} = \frac{\epsilon/m}{1+\sum_{i=1}^n 1/z_i}$, we obtain:

$$CS(Dereg.) = CS(p^*) = \frac{\epsilon^2}{2m}\left[1 - \frac{1}{1+\sum_{i=1}^n \widetilde{\beta}_i}\right]^2$$

and

$$CS(Centr.) = CS(\widetilde{p}) = \frac{\epsilon^2}{2m}\left[1 - \frac{1}{1+\sum_{i=1}^n \frac{1}{z_i}}\right]^2$$

Similarly, the expression of the generators profit is given in (4.4):

$$\text{Profit}(p, \mathbf{q}) = \sum_{i=1}^{n} [p\ q_i - C_i(q_i)]$$

For the deregulated case, we take the equilibrium price $p^* = \frac{\epsilon/m}{1+\sum_{i=1}^{n} \widetilde{\beta}_i}$ of (4.13) and the quantities $q_i^* = \beta_i\ p^*$. This leads to:

$$\text{Profit}(Dereg.) = \text{Profit}(p^*, \boldsymbol{\beta}\ p^*) = \frac{\epsilon^2}{m} \frac{1}{(1+\sum_{i=1}^{n} \widetilde{\beta}_i)^2} \left[ \sum_{i=1}^{n} \widetilde{\beta}_i (1 - \frac{z_i}{2} \widetilde{\beta}_i) \right]$$

And for the regulated case, we use the centralized price $\widetilde{p} = \frac{\epsilon/m}{1+\sum_{i=1}^{n} 1/z_i}$ of (4.4.2) and the fact that generators price at marginal costs $\widetilde{q}_i = (C_i')^{-1}(\widetilde{p}) = \frac{\widetilde{p}}{c_i}$. Substituting, we get:

$$\text{Profit}(Centr.) = \text{Profit}(\widetilde{p}, \frac{1}{\mathbf{c}}\ \widetilde{p}) = \frac{\epsilon^2}{2m} \frac{1}{(1+\sum_{i=1}^{n} \frac{1}{z_i})^2} \left[ \sum_{i=1}^{n} \frac{1}{z_i} \right]$$

The social welfare (total surplus) is then simply the sum of the consumer surplus and the generators profit:

$$TS(Dereg.) = CS(Dereg.) + \text{Profit}(Dereg.)$$

and

$$TS(Centr.) = CS(Centr.) + \text{Profit}(Centr.)$$

We now prove the bounds on the supply bid coefficients $\beta_i$'s used in evaluating the performance ratios.

**Lemma C.1.** *The slopes of the supply function bids $\widetilde{\beta}_i$'s can be bounded above and below with increasing degrees of accuracy. In particular, the following two bounds hold:*

$$\frac{1}{1+z_i} \leq \widetilde{\beta}_i \leq \frac{1}{z_i} \quad \textit{and} \quad \frac{1+\sum_{j\neq i} \frac{1}{1+z_j}}{1+z_i(1+\sum_{j\neq i} \frac{1}{1+z_j})} \leq \widetilde{\beta}_i \leq \frac{1+\sum j \neq i \frac{1}{z_j}}{1+z_i \sum_{j\neq i} \frac{1}{z_j})}$$

*Proof.* Denoting by $X_i \triangleq \sum_{j \neq i} \widetilde{\beta}_j$ and making use of equation (4.12), we can rewrite $\widetilde{\beta}_i$ as a function of $z_i$ and the $\widetilde{\beta}_j$'s:

$$\widetilde{\beta}_i = \frac{1 + \sum_{j \neq i} \widetilde{\beta}_j}{1 + z_i(1 + \sum_{j \neq i} \widetilde{\beta}_j)} = \frac{1 + X_i}{1 + z_i(1 + X_i)} \tag{C.13}$$

It turns out that the derivative of the above ratio with respect to $X_i$ is non-negative:

$$\frac{\partial \widetilde{\beta}_i}{\partial X_i} = \frac{1}{[1 + z_i(1 + X_i)]^2} \geq 0$$

As a first order approximation, we can plug the naive bound: $0 \leq X_i \leq \infty$ into (C.13). By monotonicity of $\widetilde{\beta}_i$ with respect to $X_i$, we obtain:

$$\frac{1}{1 + z_i} \leq \widetilde{\beta}_i \leq \frac{1}{z_i}$$

This in turns gives rise to the bound: $\sum_{j \neq i} \frac{1}{1+z_j} \leq X_i \leq \sum_{j \neq i} \frac{1}{z_j}$. Plugging this back into (C.13), we get:

$$\frac{1 + \sum_{j \neq i} \frac{1}{1+z_j}}{1 + z_i(1 + \sum_{j \neq i} \frac{1}{1+z_j})} \leq \widetilde{\beta}_i \leq \frac{1 + \sum_{j \neq i} \frac{1}{z_j}}{1 + z_i \sum_{j \neq i} \frac{1}{z_j})}$$

The procedure could be repeated again to improve accuracy but at the cost of increasingly complex expressions. □

Finally, we derive here bounds on the performance ratios under the deregulated setting versus the centralized setting.

**Theorem C.1.** *The generators are always better off in terms of total profit under free competition than under central coordination. Moreover the ratio of generators' profit under free competition, compared to the centrally coordinated setting, can be* *upper bounded by a function of the $z_i$'s. That is,*

$$1 \leq \frac{\textit{Profit(Dereg.)}}{\textit{Profit(Centr.)}} \leq \frac{\frac{1}{(1+\sum_{i=1}^{n}\frac{1}{1+z_i})^2}\sum_{i=1}^{n}\frac{2+z_i}{(1+z_i)^2}}{\frac{1}{(1+\sum_{i=1}^{n}\frac{1}{z_i})^2}\left[\sum_{i=1}^{n}\frac{1}{z_i}\right]} \tag{C.14}$$

*The ratio of generators' profit can be further upper bounded by a function of the minimum production cost of the generators $z_{min}$ and a measure of the equivalent number of generators $r$. That is,*

$$\frac{\textit{Profit(Dereg.)}}{\textit{Profit(Centr.)}} \leq \frac{(2+z_{min})(r+z_{min})^2}{z_{min}(1+r+z_{min})^2} \tag{C.15}$$

*Proof.* We denote by $\text{Profit}(\widetilde{\beta}_1,\ldots,\widetilde{\beta}_n) = \frac{\epsilon^2}{m}\frac{1}{(1+\sum_{i=1}^{n}\widetilde{\beta}_i)^2}\left[\sum_{i=1}^{n}\widetilde{\beta}_i(1-\frac{z_i}{2}\widetilde{\beta}_i)\right]$, the profit of generators under the deregulated scenario as a function of the $\widetilde{\beta}_i$'s. We study the monotonicity of Profit(.) with respect to $\widetilde{\beta}_i$. Taking the partial derivative of Profit(.), we get:

$$\begin{aligned}
\frac{\partial \text{Profit}}{\partial \widetilde{\beta}_i}(\widetilde{\boldsymbol{\beta}}) &= \frac{\epsilon^2}{m}\;\frac{(1-z_i\widetilde{\beta}_i)(1+\sum_{j=1}^{n}\widetilde{\beta}_j) - 2\sum_{j=1}^{n}\widetilde{\beta}_j(1-\frac{z_j}{2}\widetilde{\beta}_j)}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \\
&= \frac{\epsilon^2}{m}\;\frac{(1-z_i\widetilde{\beta}_i) + (1-z_i\widetilde{\beta}_i)\sum_{j\neq i}\widetilde{\beta}_j + (1-z_i\widetilde{\beta}_i)\;\widetilde{\beta}_i - \sum_{j=1}^{n}\widetilde{\beta}_j - \sum_{j=1}^{n}\widetilde{\beta}_j(1-z_j\widetilde{\beta}_j)}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \\
&= \frac{\epsilon^2}{m}\;\frac{(1-z_i\widetilde{\beta}_i) + (1-z_i\widetilde{\beta}_i)\sum_{j\neq i}\widetilde{\beta}_j - \sum_{j=1}^{n}\widetilde{\beta}_j - \sum_{j\neq i}\widetilde{\beta}_j(1-z_j\widetilde{\beta}_j)}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \\
&\leq \frac{\epsilon^2}{m}\;\frac{\frac{1}{1+z_{min}} + \frac{1}{1+z_{min}}\sum_{j\neq i}\widetilde{\beta}_j - \sum_{j=1}^{n}\widetilde{\beta}_j - \sum_{j\neq i}\frac{z_{min}}{z_j}\frac{1}{1+z_{min}} + \sum_{j\neq i}\frac{z_{min}}{1+z_{min}}\widetilde{\beta}_j}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \\
&\quad (\text{for } \widetilde{\beta}_i \geq \frac{1}{z_i/z_{min}+z_i}) \\
&= \frac{\epsilon^2}{m}\;\frac{\frac{1}{1+z_{min}} - \widetilde{\beta}_i - \sum_{j\neq i}\frac{z_{min}}{z_j}\frac{1}{1+z_{min}}}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \leq \frac{\epsilon^2}{m}\;\frac{\frac{1-r}{1+z_{min}}}{(1+\sum_{j=1}^{n}\widetilde{\beta}_j)^3} \qquad (\text{for } \widetilde{\beta}_i \geq \frac{1}{z_i/z_{min}+z_i}) \\
&\leq 0 \qquad (\text{since } r \geq 1)
\end{aligned}$$

Since Profit(.) is a decreasing function of $\widetilde{\beta}_i$, we can bound the generators' profit ratio using the bounds on $\widetilde{\beta}_i$ established in Lemma 4.1. Using the bound $\frac{1}{1+z_i} \leq \widetilde{\beta}_i \leq \frac{1}{z_i}$, we obtain (C.14) and with the looser bound $\{\widetilde{\beta}_j \geq \frac{1}{z_j/z_{min}+z_j} | \forall j\}$, we obtain (C.15). □

**Theorem C.2.** *The ratio of social welfare under free competition compared to the social welfare under central coordination can be upper and lower bounded by:*

$$\frac{\left[\frac{\sum_{i=1}^{n} 1/(1+z_i)}{1+\sum_{i=1}^{n} 1/(1+z_i)}\right]^2 + \frac{\sum_{i=1}^{n} 1/z_i}{(1+\sum_{i=1}^{n} 1/z_i)^2}}{1 - \frac{1}{1+\sum_{i=1}^{n} 1/z_i}} \leq \frac{TS(Dereg.)}{TS(Centr.)} \leq 1 \tag{C.16}$$

*The ratio of social welfare can be further lower bounded by a function of the minimum production cost of the generators $z_{min}$ and the measure $r$ of equivalent number of generators. That is,*

$$\frac{z_{min}}{r + z_{min}} + \frac{r\ (r + z_{min})}{(1 + r + z_{min})^2} \leq \frac{TS(Dereg.)}{TS(Centr.)} \leq 1 \tag{C.17}$$

*Proof.* The social welfare is the sum of the consumer surplus and the generators' profit. Decomposing, we get:

$$TS(Dereg.) = \frac{\epsilon^2}{2m} \left\{ \underbrace{\left[1 - \frac{1}{1 + \sum_{i=1}^{n} \widetilde{\beta}_i}\right]^2}_{\text{Increasing function of } \widetilde{\beta}_i} + \underbrace{\frac{2}{(1 + \sum_{i=1}^{n} \widetilde{\beta}_i)^2} \left[\sum_{i=1}^{n} \widetilde{\beta}_i (1 - \frac{z_i}{2} \widetilde{\beta}_i)\right]}_{\text{Decreasing function of } \widetilde{\beta}_i} \right\}$$

We thus replace the $\widetilde{\beta}_i$'s by the appropriate bounds of Lemma 4.1 to obtain the lower bound (C.16). The lower bound in (C.17) is derived using the looser bound $\{\widetilde{\beta}_j \geq \frac{1}{z_j/z_{min}+z_j} | \forall j\}$. Since under the centrally coordinated setting, electricity is dispatched by the system operator precisely to maximize social welfare, free competition cannot achieve better social welfare. The social welfare ratio is thus upper bounded by 1. □

# Bibliography


[1] E. Adida and G. Perakis. Dynamic pricing and inventory control: Uncertainty and competition. *Operations Research*, pages 1–14, 2009. Articles in Advance.

[2] E. J. Anderson and A. B. Philpott. Using supply functions for offering generation into an electricity market. *Oper. Res.*, 50, May 2002.

[3] Edward J. Anderson and Xinmin Hu. Finding supply function equilibria with asymmetric firms. *OPERATIONS RESEARCH*, 56(3):697–711, 2008.

[4] J.M. Arroyo and F.D. Galiana. Energy and reserve pricing in security and network-constrained electricity markets. *Power Systems, IEEE Transactions on*, 20(2):634 – 643, May 2005.

[5] J. Bain. Relation of profit rate to industry concentration: American manufacturing,1936-1940. *Quarterly Journal of Economics*, 65:293–324, 1951.

[6] Ross Baldick, Ryan Grant, and Edward Kahn. Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics*, pages 143–167, 2004.

[7] Ross Baldick and William Hogan. Stability of supply function equilibria implications for daily versus hourly bids in a poolco market. *Journal of Regulatory Economics*, 30(2):119–139, 2006.

[8] R. Bamon and J. Fraysśe. Existence of cournot equilibrium in large markets. *Econometrica*, 53:587–597, 1985.

[9] M. T. Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12(4):287–298, 2002.

[10] W. J. Baumol and W. E. Oates. The use of standards and prices for protection of the environment. *The Swedish Journal of Economics,*, 73:42–54, 1971.

[11] A. Bensoussan. Points de nash dans le cas de fonctionnelles quadratiques et jeux differentiels lineaires a n personnes. *SIAM J Control*, 12:460–499, 1974.

[12] F. Bernstein and A. Federgruen. Pricing and replenishment strategies in a distribution system with competing retailers. *Operations Research*, 3:409–426, January 2003.

[13] F. Bernstein and A. Federgruen. Comparative statics, strategic complements and substitutes in oligopolies. *Journal of Mathematical Economics*, 40:713–746, 2004.

[14] F. Bernstein and A. Federgruen. Decentralized supply chains with competing retailers under demand uncertainty. *Management Science*, 51:18–29, January 2005.

[15] F. Bernstein and A. Federgruen. Coordination mechanisms for supply chains under price and service competition. *MSOM*, 9:242–262, 2007.

[16] Carolyn A. Berry, Benjamin F. Hobbs, William A. Meroney, Richard P. O'Neill, and William R. Stewart. Understanding how market power can arise in network competition: a game theoretic approach. *Utilities Policy*, 8(3):139 – 158, 1999.

[17] J. Bertrand. Théorie mathématique de la richesse sociale. *Journal des Savants*, pages 499–508, 1983.

[18] Boris Blaesig and Hans-Jurgen Haubrich. Optimization of risk management problems in generation and trading planning. *Optimization in the Energy Industry*, pages 387–408, 2009.

[19] Allaz Blaise and Vila Jean-Luc. Cournot competition, forward markets and efficiency. *Journal of Economic Theory*, 59(1):1–16, February 1993.

[20] R. Bordley. Relating cross-elasticities to first choice/second choice data. *Journal of Business and Economic Statistics*, 1986.

[21] Severin Borenstein and James Bushnell. An empirical analysis of the potential for market power in california's electricity industry. *NBER Working Papers*, March 1998.

[22] F. Bouffard, F.D. Galiana, and A.J. Conejo. Market-clearing with stochastic security-part i: formulation. *Power Systems, IEEE Transactions on*, 20(4):1818 – 1826, November 2005.

[23] D.W. Bunn. Forecasting loads and prices in competitive power markets. *Proceedings of the IEEE*, 88(2):163 –169, February 2000.

[24] G. Cachon. Supply chain coordination with contracts. *Operations Research and Management Science*, pages 229–239, 2002.

[25] G.P Cachon and M.A Lariviere. Supply chain coordination with revenue-sharing contracts. *Management Science*, 51:30–44, 2005.

[26] L. M.A. Chan, P. R. Goundan, and D. Simchi-Levi. The impact of pricing and buy-back menus on supply chain performance. *Working Paper, MIT*, 2006.

[27] Chunhui Chen and O. L. Mangasarian. A class of smoothing functions for nonlinear and mixed complementarity problems. *Comput. Optim. Appl.*, 5:97–138, March 1996.

[28] F. Chen, J. K. Ryan Z. Drezner, and D. Simchi-Levi. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science*, 46:436–443, 2000.

[29] Frank Chen, Zvi Drezner, Jennifer K. Ryan, and David Simchi-Levi. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Manage. Sci.*, 46:436–443, March 2000.

[30] J. Contreras, M. Klusch, and J.B. Krawczyk. Numerical solutions to nash-cournot equilibria in coupled constraint electricity markets. *Power Systems, IEEE Transactions on*, 19(1):195 – 206, February 2004.

[31] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses. Selfish routing in capacitated networks,. *Mathematics of Operations Research*, 29:961–976, 2004.

[32] J.R. Correa, N. Figueroa, and N. Stier-Moses. Pricing with markups in industries with increasing marginal costs'. *DRO, Columbia University*, 2009.

[33] A. Cournot. Recherches sur les principes mathematics de la theorie de la richesse. *Hachettte*, 1838.

[34] K. Cowling and M. Waterson. Price-cost margins and market structure. *Economic Journal*, 43:267–274, 1976.

[35] B. Curry and K. George. Industrial concentration: A survey. *Journal of Industrial Economics*, 31:203–255, 1983.

[36] R. Dansby and R. Willig. Industry performance gradient indexes. *American Economic Review*, 69:249–260, 1979.

[37] P. Dasgupta and E. Maskin. The existence of equilibrium in discontinuous economic games. *Review of Economic Studies*, 53:1–41, 1986.

[38] Martinez de Albeniz and V. D. Simchi-Levi. Competition in the supply option market,. 2003. Working paper, MIT, .

[39] G. Debreu. A social equilibrium existence theorem. *Proc Natl Acad Sci*, 38:886–893, 1952.

[40] R. Deneckere and C. Davidson. Incentives to form coalitions with bertrand competition. *Rand Journal of Economics*, 16:473–486, 1985.

[41] Shi-Jie Deng and Wenjiang Jiang. Levy process-driven mean-reverting electricity price model: the marginal distribution analysis. *Decis. Support Syst.*, 40:483–494, October 2005.

[42] Steven P. Dirkse and Michael C. Ferris. Mcplib: A collection of nonlinear mixed complementarity problems. *Optimization Methods and Software*, 5:319–345, 1994.

[43] Steven P. Dirkse and Michael C. Ferris. The path solver: A non-monotone stabilization scheme for mixed complementarity problems. *Optimization Methods and Software*, 5:123–156, 1995.

[44] D. Encaoua and A. Jacquemin. Degree of monopoly, indices of concentration, and threat of entry. *International Economic Review*, 21:87–105, 1980.

[45] F. Facchinei and C. Kanzow. Generalized nash equilibrium problems. *4OR: A Quarterly Journal of Operations Research*, 5:173–210, 2007.

[46] Farahat and G. Perakis. A comparison of price and quantity competition in oligopolies with differentiated products. *Accepted in Operations Research.*

[47] A. Farahat and G. Perakis. Price competition among multiproduct firm. *Working Paper.*

[48] Amr Farahat and Georgia Perakis. Profit loss in differentiated oligopolies. *Operations Research Letters*, 37:43 – 46, 2009.

[49] J. Farell and C. Shapiro. Horizontal mergers: An equilibrium analysis. *American Economic Review*, 80:107–126, 1990.

[50] FERC. Notice of proposed rulemaking. revised filing requirements under part 33 of the commissions regulations. 1998.

[51] O.B. Fosso, A. Gjelsvik, A. Haugstad, B. Mo, and I. Wangensteen. Generation scheduling in a deregulated system. the norwegian case. *Power Systems, IEEE Transactions on*, 14:75 –81, February 1999.

[52] J. W. Friedman. On the strategic importance of prices versus quantities. *RAND Journal of Economics*, 19:607–622, 1988.

[53] Douglas T. Gardner and J. Scott Rogers. Planning electric power systems under demand uncertainty with different technology lead times. *Manage. Sci.*, 45:1289–1306, October 1999.

[54] B.G. Gorenstin, N.M. Campodonico, J.P. Costa, and M.V.F. Pereira. Stochastic optimization of a hydro-thermal system including network constraints. *Power Industry Computer Application Conference, 1991. Conference Proceedings*, pages 127 –133, May 1991.

[55] Mika Goto and G. Andrew Karolyi. Understanding electricity price volatility within and across markets. *Working Paper Series*, July 2004.

[56] Richard J Green and David M Newbery. Competition in the british electricity spot market. *Journal of Political Economy*, 100(5):929–53, 1992.

[57] Richard John Green. Increasing competition in the british electricity spot market. *Journal of Industrial Economics*, 44(2):205–16, 1996.

[58] G. Gross and F.D. Galiana. Short-term load forecasting. *Proceedings of the IEEE*, 75:1558 - 1573, December 1987.

[59] Guidelines for Energy Management, http://www.energystar.gov/index.cfm?fuseaction=partners_in_practice.showStory&storyID=1000100&step=IP. Retrieved April 28, 2010.

[60] Xiaolei Guo and Hai Yang. The price of anarchy of cournot oligopoly. *Internet and Network Economics*, 3828:246–257, 2005.

[61] A. Haurie and J. Krawczyk. Optimal charges on river effluent from lumped and distributed sources. *Environmental Modeling and Assessment*, 2:177189, 1997.

[62] B.E. Hobbs. Linear complementarity models of nash-cournot competition in bilateral and poolco power markets. *Power Systems, IEEE Transactions on*, 16(2):194 -202, May 2001.

[63] Benjamin F. Hobbs and J. S. Pang. Nash-cournot equilibria in electric power markets with piecewise linear demand functions and joint constraints. *OPERATIONS RESEARCH*, 55(1):113–127, 2007.

[64] B.F. Hobbs, C.B. Metzler, and J.-S. Pang. Strategic gaming analysis for electric power systems: an mpec approach. *Power Systems, IEEE Transactions on*, 15(2):638 -645, May 2000.

[65] M. Hoel and L. Karp. Taxes versus quotas for a stock pollutant. *Resource and Energy Economics*, 24:367–384, 2010.

[66] William W. Hogan. A market power model with strategic interaction in electricity networks. *The Energy Journal*, 18(4):107–142, 1997.

[67] Kenneth Holmstrm, Marcus M. Edvall, and Anders O. Gran. Tomlab. Technical report.

[68] R. A. Horn and C. R. Johnson. Matrix analysis. *Cambridge University Press*, 1985.

[69] Xinmin Hu and Daniel Ralph. Using epecs to model bilevel games in restructured electricity markets with locational prices. *OPERATIONS RESEARCH*, 55(5):809–827, 2007.

[70] Intel Corporate Responsibility Report 2009, http://www.intel.com/Assets/PDF/Policy/CSR-2009.pdf. Retrieved June 18, 2010.

[71] Wei Jing-Yuan and Yves Smeers. Spatial oligopolistic electricity models with cournot generators and regulated transmission prices. *Oper. Res.*, 47:102–112, January 1999.

[72] C. R. Johnson. Inverse m-matrices. *Linear Algebra and its Applications*, 47:195–216, 1982.

[73] L. Karp. Nonpoint source pollution taxes and excessive tax burden. *Environmental & Resource Economics*, 31:229–251, 2005.

[74] Daniel Kirschen and Goran Strbac. Fundamentals of power system economics. 2004.

[75] Paul D Klemperer and Margaret A Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243–77, November 1989.

[76] J. Kluberg and G. Perakis. Generalized quantity competition for multiple products and loss of efficiency. *Forthcoming, Operations Research*, 2010.

[77] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. *Proceedings of the 16th annual conference on Theoretical aspects of computer science*, 1563:404–413, 1999.

[78] J. Krawczyk. Coupled constraint nash equilibria in environmental games. *Resource and Energy Economics*, 27:157181, 2005.

[79] Jacek Krawczyk and James Zuccollo. Nira-3: An improved matlab package for finding nash equilibria in infinite games. MPRA Paper 1119, University Library of Munich, Germany, December 2006.

[80] D. Kreps and J. Scheinkman. Cournot precommitment and bertrand competition yield cournot outcomes. *Bell Journal of Economics*, 14:326–337, 1983.

[81] M. Lariviere. Supply chain contracting and coordination with stochastic demand. *Quantitative Models for Supply Chain Management*, pages 233–268, 1999.

[82] M. Lariviere and E. Porteus. Selling to the newsvendor: An analysis of price only contracts. *Manufacturing and Service Operations Management*, 3:293–305, 2001.

[83] Z.-Q. Luo, J.-S. Pang, and D. Ralph. Mathematical programs with equilibrium constraints. 1996.

[84] A. Maiorano, Y.H. Song, and M. Trovato. Dynamics of non-collusive oligopolistic electricity markets. 2:838 –844, 2000.

[85] O. Mancino and G. Stampacchia. Convex programming and variational inequalities. *Journal of Optimization Theory and Application*, 9:3–23, 1972.

[86] M. McManus. Equilibrium, number and size in cournot oligopoly. *Yorkshire Bulletin of Social and Economic Research*, 16:68–75, 1964.

[87] W. D. Montgomery. Markets in licenses and efficient pollution control programs. *Journal of Economic Theory*, 5:395–418, 1972.

[88] A. Monticelli, M. V. F. Pereira, and S. Granville. Security-constrained optimal power flow with post-contingency corrective rescheduling. *Power Systems, IEEE Transactions on*, 2(1):175 –180, February 1987.

[89] F. H. Murphy, H. D. Sherali, and A. L. Soyster. A mathematical programming approach for determining oligopolistic market equilibrium. *Mathematical Programming*, 24:92–106, 1982.

[90] Frederic H. Murphy and Yves Smeers. Generation capacity expansion in imperfectly competitive restructured electricity markets. *OPERATIONS RESEARCH*, 53(4):646–661, 2005.

[91] A. Nagurney and K. K. Dhanda. Variational inequalities for marketable pollution permits with technological investment opportunities: The case of oligopolistic markets. *Mathematical Comput. Modelling*, 26:1–25, 1997.

[92] J. Nash. Non-cooperative games,. *Annals of Mathematics*, 54:128–140, 1951.

[93] W. Novshek. On the existence of cournot equilibrium. *Review of Economic Studies*, 52:85–98, 1985.

[94] A. M. Ostrowski. Note on bounds for determinants with dominant principal diagonal. volume 3, pages 26–30, 1952.

[95] A.L. Ott. Experience with pjm market operation, system design, and implementation. *Power Systems, IEEE Transactions on*, 18(2):528 – 534, May 2003.

[96] V. Padmanabhan and I. P. L. Png. *Management of Environmental Activities, Company Website.* Panasonic, http://panasonic-electric-works.net/csr/environment /management/promotion/index.html, 1995. Retrieved June 23, 2010.

[97] V. Padmanabhan and I. P.L. Png. Manufacturers returns policies and retail competition. *Sloan Marketing Science*, 16:81–94, 1995.

[98] J.-S. Pang and M. Fukushima. Quasi-variational inequalities, generalized nash equilibria, and multi-leader-follower games. *Computational Management Science*, 2:21–56, 2005.

[99] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang. Distributed power allocation with rate constraints in gaussian parallel interference channels. *IEEE Transactions on Information Theory*, 2008.

[100] I. W.H. Parry. Environmental taxes and quotas in the presence of distorting taxes in factor markets. *Resource and Energy Economics*, 19:203–220, 1997.

[101] B. Pasternack. Optimal pricing and returns policies for perishable commodities. *Marketing Science*, 4:166–176, 1985.

[102] G. Perakis. The price of anarchy under nonlinear and asymmetric costs,. *Mathematics of Operations Research*, 32:614–628, 2007.

[103] G. Perakis and G. Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts,. *Management Science*, 53:1249–1268, 2007.

[104] Georgia Perakis and Guillaume Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. *MANAGEMENT SCIENCE*, 53(8):1249–1268, 2007.

[105] M. V. F. Pereira and L. M. V. G. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52:359–375, 1991.

[106] A. B. Philpott, M. Craddock, and H. Waterer. Hydro-electric unit commitment subject to uncertain demand. *European Journal of Operational Research*, 125(2):410–424, September 2000.

[107] Power Systems Test Case Archive, http://www.ee.washington.edu/research/pstca /pf30/pg_tca30bus.htm. *30 Bus Power Flow Test Case.*

[108] Proctor and Gamble, http://www.environmentalleader.com/2009/10/20/pg-cuts-co2-emissions-and-water-use-by-52-since-2002.

[109] H. Marvel R. Deneckere and J. Peck. Demand uncertainty and price maintenance: Markdowns as destructive competition. *Amer. Econom. Rev.*, 87:619–641, 1997.

[110] R. S. Varga R. Nabben. A linear algebra proof that the inverse of a strictly ultrametric marix is a strictly diagonally dominant stieltjes matrix. *SIAM Journal of Matrix Analysis Applications*, 15:107–113, January 1994.

[111] J. Roberts and H. Sonnenschein. On the existence of cournot equilibrium without concave profit functions. *Journal of Economic Theory*, 13:112–117, 1976.

[112] R. M. Rockafellar. Convex functions, monotone operators, and variational inequalities. *Theory and Applications of Monotone Operators, Proceedings of the NATO Advanced Study Institute, Venice, Italy*, 1968. (Edizioni Oderisi, Gubbio, Italy, 1968).

[113] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica: Journal of the Econometric Society*, 33:520–534, 1965.

[114] T. Roughgarden and E. Tardos. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games and Economic Behavior*, 47, 2004.

[115] Tim Roughgarden and Éva Tardos. How bad is selfish routing? *J. ACM*, 49:236–259, March 2002.

[116] Aleksandr Rudkevich. Supply function equilibrium in power markets: Learning all the way. *TCA Technical Paper*, pages 1299–1702, 1999.

[117] S. Salant, S. Switzer, and R. Reynolds. Losses due to merger: The effects of an exogenous change in industry structure on cournot-nash equilibrium. *Quarterly Journal of Economics*, 48:185–200, 1983. .

[118] R. Schmalensee. Inter-industry studies of structure and performance. *Handbook of Industrial Organization*, 1986.

[119] M. Shubik and R. Levitan. Market structure and behavior. *Cambridge: Harvard University Press*, 1980. .

[120] N. Singh and X. Vives. Price and quantity competition in a differentiated duopoly. *Rand Journal of Economics*, 15:546–554, 1984.

[121] Joseph J. Spengler. Vertical integration and antitrust policy. *Journal of Political Economy*, 58:347, 1950.

[122] G. Stigler. A theory of oligopoly. *Journal of Political Economy*, 72:225–258, 1964.

[123] J. Sun. A convergence analysis for a convex version of dikin's algorithm. *Annals of Operations Research*, 62:357–373, 1996.

[124] F. Szidarowsky and S. Yakowitz. A new proof of the existence and uniqueness of the cournot equilibrium. *International Economic Review*, 18:787–789, 1977.

[125] F. Szidarowszky and K. Okuguchi. A linear oligopoly model with adaptive expectations: Stability reconsidered. *Journal of Economics*, 48:79–82, 1988.

[126] S. Takriti, J. R. Birge, and E. Long. A stochastic model for the unit commitment problem. *Power Systems, IEEE Transactions on*, (3):1497–1508, 1996.

[127] M. Tidball and G. Zaccour. An environmental game with coupling constraints. *Environmental Modeling and Assessment*, 10:153158, 2005.

[128] J. Tirole. The theory of industrial organization. *MIT Press*, 1988.

[129] A. Tsay, S. Nahmias, and N. Agrawal. Modeling supply chain contracts: A review. *Quantitative Models for Supply Chain Management*, page 299336, 1999.

[130] Chung-Li Tseng and Graydon Barz. Short-term generation asset valuation. *Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences-Volume 3 - Volume 3*, pages 3028–, 1999.

[131] Stephen J. Turnbull. Choosing duopoly solutions by consistent conjectures and by uncertainty. *Economics Letters*, 13(2-3):253 – 258, 1983.

[132] U.S. Energy Information Administration, http://www.eia.gov/electricity/data.cfm#summary. *Electricity Data.*

[133] US Wireless cellular data, http://www.slideshare.net/chetansharma/us-wireless-market-q4-2008-and-2008-update-mar-2009-chetan-sharma-consulting.

[134] X. Vives. Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 1990.

[135] Xavier Vives. *Oligopoly Pricing: Old Ideas and New Tools.* The MIT Press, June 2001.

[136] Xavier Vives. Strategic supply function competition with private information. 2009.

[137] M. L. Weitzman. Prices versus quantities. *Review of Economic Studies*, 41:477–491, 1974. .

[138] S. Wellisz. On external diseconomies and the government-assisted invisible hand. *Economica*, 31:345–362, 1964.

[139] S. Wong and J.D. Fuller. Pricing energy and reserves using stochastic optimization in an alternative electricity market. *Power Systems, IEEE Transactions on*, 22(2):631 –638, May 2007.

[140] Allen J. Wood and Bruce F. Wollenberg. *Power generation, operation, and control / Allen J. Wood, Bruce F. Wollenberg.* Wiley, New York :, 1984.

[141] C. Wright. Some aspects of the use of corrective taxes for controlling air pollution emissions. *Journal of Natural Resources*, 9:63–84, 1969.

[142] Alexander Zaslavski, Richard Byrd, Jorge Nocedal, and Richard Waltz. Knitro: An integrated package for nonlinear optimization. 83:35–59, 2006.